

**Universidad Autónoma de Madrid**  
**Escuela Politécnica Superior**  
**Máster en Ingeniería Informática y Telecomunicaciones**



***Detección de emociones en textos y su  
aplicación a la categorización de obras y  
autores***

**Trabajo Fin de Máster**

**Linda E. Barros Kaiser**

**Noviembre 2013**







# ÍNDICE

---

<b>1</b>	<b><i>Introducción y objetivos</i></b>	<b>1</b>
1.1	Introducción	1
1.2	Objetivos	2
1.3	Organización del trabajo	4
<b>2</b>	<b><i>Detección de Emociones</i></b>	<b>5</b>
2.1	Detección de emociones en textos	5
2.2	Detección de emociones en el habla	8
2.3	Detección de emociones en expresiones faciales y corporales	9
2.4	Detección combinada de emociones (multimodal)	10
<b>3</b>	<b><i>Proceso de construcción de los conjuntos emocionales</i></b>	<b>13</b>
3.1	Construyendo los conjuntos emocionales	14
3.1.1	Procedimiento de identificación de palabras que denotan emoción	14
3.1.2	Construcción de los conjuntos emocionales en español	15
3.2	Validación del enfoque propuesto	16
3.2.1	ANEW y su adaptación al español	16
3.2.2	Comparando el enfoque propuesto en español con la adaptación española de ANEW	17
3.3	Clasificación no supervisada de las palabras del conjunto de emociones a partir del contenido de S-ANEW	21
3.4	Discusión	22
<b>4</b>	<b><i>Análisis de la categorización automática de autores y obras literarias</i></b>	<b>23</b>
4.1	Categorización emocional de perfiles emocionales de autores y conversaciones	23
4.1.1	Autores de cuentos infantiles	23
4.1.2	G. A. Bécquer y M. Benedetti	25
4.1.2.1	Gustavo Adolfo Bécquer	25
4.1.2.2	Mario Benedetti	28
4.1.2.3	Discusión de las obras de Bécquer y Benedetti	31
4.1.3	Categorización de conversaciones en un contexto específico	31

4.2	Categorización emocional, no supervisada, de los poemas de Francisco Quevedo..	33
4.2.1	Categorización de los datos originales .....	34
4.2.2	Resultados.....	35
4.2.2.1	Dataset original: 185 instancias y 4 clases .....	35
4.2.2.2	Resampling: 740 instancias y 4 clases .....	37
4.2.2.3	Uniando y eliminando clases .....	39
4.2.2.4	Otras técnicas de aprendizaje.....	39
4.2.3	Discusión .....	40
<b>5</b>	<b><i>Conclusiones y trabajos futuros.....</i></b>	<b>43</b>
5.1	Conclusiones.....	43
5.2	Trabajos futuros.....	46
	<b><i>Referencias.....</i></b>	<b><i>i</i></b>
	<b><i>Apéndice 1: Construyendo los conjuntos emocionales en inglés .....</i></b>	<b><i>ix</i></b>
	<b><i>Apéndice 2: Construyendo los conjuntos emocionales en español.....</i></b>	<b><i>xi</i></b>
	<b><i>Apéndice 3: Listado de obras de G. A. Bécquer y M. Benedetti.....</i></b>	<b><i>xvii</i></b>

# Índice de Figuras

---

<i>Figura 3.1. SAM utilizado en S-ANEW. Escalas de autoevaluación para medir la: Valencia, Excitación y Dominio .....</i>	<i>17</i>
<i>Figura 3.2. Valencia vs. Excitación.....</i>	<i>18</i>
<i>Figura 3.3. Valencia vs. Dominio .....</i>	<i>19</i>
<i>Figura 3.4. Excitación vs. Dominio .....</i>	<i>19</i>
<i>Figura 3.5. Valencia vs. Excitación.....</i>	<i>20</i>
<i>Figura 3.6. Valencia vs. Dominio .....</i>	<i>20</i>
<i>Figura 3.7. Excitación vs. Dominio .....</i>	<i>21</i>
<i>Figura 4.1. Cuentos infantiles.....</i>	<i>24</i>
<i>Figura 4.2. Proporción de emociones en 22 cuentos de Bécquer.....</i>	<i>25</i>
<i>Figura 4.3. Cuentos de G. A. Bécquer.....</i>	<i>26</i>
<i>Figura 4.4. Agrupación de cuentos de Bécquer con la misma proporción emocional .....</i>	<i>26</i>
<i>Figura 4.5. Proporción de emociones en 15 rimas de Bécquer .....</i>	<i>27</i>
<i>Figura 4.6. Rimad de G. A. Bécquer.....</i>	<i>27</i>
<i>Figura 4.7. Agrupación de rimas de Bécquer con la misma proporción emocional .....</i>	<i>28</i>
<i>Figura 4.8. Proporción de emociones en 20 cuentos de Benedetti.....</i>	<i>28</i>
<i>Figura 4.9. Cuentos de Mario Benedetti .....</i>	<i>29</i>
<i>Figura 4.10. Agrupación de cuentos de Benedetti con la misma proporción emocional .....</i>	<i>29</i>
<i>Figura 4.11. Proporción de emociones en 20 poemas de Benedetti.....</i>	<i>30</i>
<i>Figura 4.12. Poemas de Mario Benedetti.....</i>	<i>30</i>
<i>Figura 4.13. Agrupación de poemas de Benedetti con la misma proporción emocional .....</i>	<i>31</i>

<b>Figura 4.14. Porcentaje emocional del intercambio de e-mails.....</b>	<b>32</b>
<b>Figura 4.15. Análisis emocional del intercambio de e-mails .....</b>	<b>32</b>
<b>Figura 4.16. Árbol de clasificación con el dataset original: 185 instancias, 4 clases .....</b>	<b>35</b>
<b>Figura 4.17. Sobreajuste del Resampling: 740 instancias y 4 clases (a 16 hojas como instancias mínimas).....</b>	<b>38</b>



# Índice de Tablas

---

<i>Tabla 4.1. Valores máximos de cada emoción.....</i>	<i>31</i>
<i>Tabla 4.2. Distribución original de las clases.....</i>	<i>36</i>
<i>Tabla 4.3. Matriz de confusión para el Clasificador 4.2.2.1 .....</i>	<i>36</i>
<i>Tabla 4.4. Valores-F para el clasificador 4.2.2.1.....</i>	<i>36</i>
<i>Tabla 4.5. Distribución de las 4 Clases con Resampling.....</i>	<i>37</i>
<i>Tabla 4.6. Matriz de confusión para el Clasificador 4.2.2.2 .....</i>	<i>37</i>
<i>Tabla 4.7. Matriz de confusión para el sobreajuste del Clasificador 4.2.2.2..</i>	<i>38</i>
<i>Tabla 4.8. Separando las clases por pares.....</i>	<i>39</i>
<i>Tabla 5.1. Patrones principales para cada clase.....</i>	<i>41</i>
<i>Tabla Ap.2. Conjunto de palabras en español para “Joy y Joyful” .....</i>	<i>xi</i>
<i>Tabla Ap.3. Conjunto de palabras en español para “Sadness y Sad” .....</i>	<i>xii</i>
<i>Tabla Ap.4. Conjunto de palabras en español para “Anger y Angry” .....</i>	<i>xiii</i>
<i>Tabla Ap.5. Conjunto de palabras en español para “Fear y Afraid” .....</i>	<i>xv</i>
<i>Tabla Ap.6. Listado de 22 cuentos de Bécquer .....</i>	<i>xvii</i>
<i>Tabla Ap.7. Listado de 15 rimas de Bécquer.....</i>	<i>xvii</i>
<i>Tabla Ap.8. Listado de 20 cuentos de Benedetti.....</i>	<i>xviii</i>
<i>Tabla Ap.9. Listado de 20 poemas de Benedetti .....</i>	<i>xviii</i>



# 1 Introducción y objetivos

---

## 1.1 Introducción

Las emociones son un componente intrínseco del ser humano que afecta a todos los aspectos de su conducta y, en particular, a aquellos relacionados con la comunicación. Las personas estamos constantemente transmitiendo emociones de una forma u otra; por esta razón, el análisis de emociones se ha convertido en un área de investigación multidisciplinar de interés creciente.

En este contexto, dentro del conjunto de las llamadas “emociones básicas”, Paul Ekman [38] encontró evidencias que respaldan la universalidad de las emociones que, según este trabajo, están representadas por: *alegría*, *tristeza*, *furia o pasión*, *miedo*, *disgusto* y *sorpresa* (identificadas por las palabras “joy”, “sadness”, “anger”, “fear”, “disgust” y “surprise”, respectivamente). A partir del trabajo de Ekman [38], Azcarate et al. [8] proponen un nuevo grupo restringido de emociones, sosteniendo que las emociones de *disgusto* y *sorpresa* no pertenecen a las llamadas “emociones básicas”, porque pueden ser fácilmente asociadas con las emociones de *furia-pasión* y *miedo*, respectivamente. Por esta razón sugieren un nuevo conjunto de emociones básicas, subconjunto del anterior: *alegría*, *furia o pasión*, *tristeza* y *miedo* (“joy”, “anger”, “sadness” y “fear”). Zinck y Newen [92] coinciden con esta aproximación, y exponen que estas cuatro emociones están directamente relacionadas con los “desafíos fundamentales” de la vida cotidiana, esto es; la seguridad en uno mismo (auto-eficiencia) y la aceptación social producen *felicidad*, el peligro o la zozobra conducen al *miedo*, la separación de condiciones favorables o la pérdida de un ser querido producen *tristeza* o por último el sentimiento de frustración sobre cualquier expectativa conduce a la *furia*, y también podríamos pensar que se asocia, de algún modo, con la *pasión*.

Por otra parte, respecto al lenguaje, éste, además de poder comunicarnos, nos ha dado la posibilidad de manifestar la estética de los sentimientos y las experiencias vividas o ficticias a través de la palabra. Con este propósito nació la literatura y, con ella, los géneros literarios: lírico, narrativo y dramático. Cada uno cumple una función concreta dentro de la literatura: el género lírico enfatiza las emociones y sensaciones del escritor, mientras los géneros narrativo y dramático, al margen de sus diferencias, tienen el propósito común de relatar tanto historias como vivencias reales o ficticias.

La detección y análisis de emociones en textos es un área con mucho potencial que se encuentra en constante crecimiento. Hasta la fecha se pueden encontrar trabajos muy interesantes en los que se aplica esta forma de analizar las emociones para medir la felicidad nacional [36], el nivel de aceptación acerca de un evento o producto [1], analizar canciones [51], discursos [35] o noticias [12]. Todo esto aplicado, en gran parte, sobre plataformas web de acceso público. Entre

ellas, las más utilizadas son las redes sociales, los blogs, sitios que almacenan y publican información o sitios de recomendación y opinión de productos. Igualmente el análisis emocional puede aplicarse sobre cualquier tipo de texto que se encuentre en formato digital.

También, además de utilizar la detección de emociones para analizar y medir cuantitativamente la presencia de emociones en textos aislados, se puede pensar en su utilidad como mecanismo clasificador. Esta posibilidad permitiría la identificación de patrones emocionales de diverso tipo. Un ejemplo de ello podría ser la búsqueda de patrones emocionales en la categorización realizada a poemas de un escritor por medio del análisis emocional de cada categoría y así encontrar el patrón emocional asignado a dichas categorías.

## 1.2 Objetivos

El objetivo principal de este trabajo es explorar las posibilidades existentes para la categorización automática de textos en función de su contenido emocional.

Los pasos principales que se realizarán en este trabajo son los siguientes:

En primer lugar, se **construirá una serie de conjuntos emocionales basados en el marco de las emociones básicas**. El enfoque presentado en este trabajo está basado en la recuperación de información a través de un procedimiento de traducción de inglés a español (*cross-linguistic information retrieval*). El punto de partida será tomar las cuatro emociones básicas antes mencionadas y, a continuación, proceder a llenar estos cuatro conjuntos con palabras que, de alguna forma u otra, evoquen a cada una de dichas emociones. Al llenar el conjunto de *anger*, éste debería dar cabida a las palabras que evocan a dicha emoción y así, igualmente, con cada una de las emociones.

El procedimiento para construir cada uno de los conjuntos requiere de varios pasos importantes, para los que partimos de nuestras cuatro emociones básicas escritas en inglés: *joy*, *sadness*, *anger* y *fear*. A partir de ellas, el primer paso será buscar en inglés los sinónimos<sup>1</sup> de cada una de las emociones, luego los adjetivos y, por último, otras palabras sugeridas por el diccionario. Posteriormente se examinará la traducción contextual de cada una de las palabras al español, agregando algunas palabras sugeridas a través de la traducción.

A continuación, se **procederá a validar los conjuntos emocionales** obtenidos. Con el fin de validar nuestra propuesta, se comparará la forma en que nuestra propuesta y S-ANEW [69] (uno de los más conocidos diccionarios “emocionales”) clasifican las palabras de sus conjuntos respectivos, y de qué tipo son las palabras contenidas en ambos. Por medio de este procedimiento se intenta averiguar si las palabras contenidas en los cuatro conjuntos

---

<sup>1</sup> Enero del 2012, <http://bab.la>, Sinónimos obtenidos de © Princeton University

emocionales comparten las mismas características que las contenidas en S-ANEW, o en otras palabras, poder probar si ambos enfoques son consistentes.

Por último, se **aplicará nuestra propuesta para realizar la categorización automática de autores y obras literarias**. Una vez creadas las bases para el reconocimiento sistemático de las emociones, el paso final consistirá en establecer algunos escenarios donde pueda ser aplicado nuestro sistema de reconocimiento de emociones. Para poder llevar a cabo este objetivo, se examinará cómo el análisis emocional a través de nuestra propuesta nos puede proveer de las pautas necesarias para poder realizar la categorización emocional de trabajos literarios y conversaciones informales desde dos puntos de vista:

- Por una parte, se intentará **categorizar perfiles por medio del análisis emocional** en tres escenarios: dos de ellos relacionados a la búsqueda de perfiles emocionales sobre cuentos infantiles y el estudio de las obras de dos autores famosos en la lengua hispana (Bécquer y Benedetti) y, el tercero, relacionado con la búsqueda de perfiles emocionales en conversaciones realizadas por correo electrónico.
- Luego, se buscará **realizar una clasificación emocional no supervisada de una serie de poemas de Francisco Quevedo** por medio del análisis emocional.

Asociadas a los objetivos anteriores subyacen algunas hipótesis e interrogantes. La **primera hipótesis** se plantea de la siguiente manera: *Al analizar emocionalmente cuentos infantiles de cinco autores, los rangos emocionales tendrían que ser parecidos, debido a que cuando se habla de cuentos infantiles se asume que la emoción dominante tendría que ser la felicidad.*

Sobre la premisa de la semejanza en la escritura se plantea la **segunda hipótesis**: *Si se analizan los trabajos de dos escritores famosos tanto en género lírico como en el narrativo, tendríamos resultados emocionales muy parecidos en cada uno de los géneros.*

Además de estas dos hipótesis, se nos plantean dos interrogantes que marcan los objetivos de los planteamientos restantes: *¿El sistema de reconociendo de emociones sería capaz de reconocer las diferencias en los sentimientos transmitidos en dos grupos de conversaciones, realizadas en un contexto específico?*

Y por último, *¿Si se toman los poemas Francisco de Quevedo y la clasificación manual de los mismos, el sistema de reconocimiento propuesto sería capaz de realizar dicha clasificación automáticamente?*

Parte de los resultados obtenidos en el desarrollo de estos objetivos han sido publicados en los siguientes trabajos:

- L. **Barros**, P. Rodríguez and A. Ortigosa, "Emotion recognition in texts for user model augmenting," in Proceedings of the 13th International Conference on Interacción Persona-Ordenador, Elche, Spain, 2012, pp. 45:1-45:6.
- L. **Barros**, P. Rodríguez and A. Ortigosa. Automatic Classification of Literature Pieces by Emotion Detection. A Study on Quevedo's Poetry. In Proceedings of the fifth biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII 2013). In Geneva Switzerland, 2013, pages 6. In press.

### ***1.3 Organización del trabajo***

El trabajo que se presenta a continuación está distribuido en varios apartados, de la siguiente manera:

A continuación, en el **apartado dos**, se incluye una revisión del estado del arte en lo que se refiere a la detección no supervisada de emociones, estudiando los diferentes enfoques en que puede ser planteado el reconocimiento de las mismas, como son el reconocimiento de emociones en los gestos faciales y corporales, en el habla y principalmente en las emociones manifestadas en la escritura.

En el **tercer apartado** se presentan la metodología y la validación de la clasificación a través de su comparación con la base de datos emocional S-ANEW [69]. Luego, en el **cuarto apartado**, se realiza la validación de la clasificación, aplicando nuestra propuesta de conjuntos emocionales con el fin de explorar las posibilidades existentes para realizar la categorización automática de textos en función de su contenido emocional.

Las conclusiones y trabajos futuro relacionados con esta propuesta se encuentran en el **quinto apartado**. Por último, tras las **referencias**, se incluyen algunos **apéndices** que pueden resultar de utilidad.

## 2 Detección de Emociones

---

Durante las últimas décadas se han desarrollado numerosas estudios con el objetivo de medir, evaluar o identificar emociones, tanto de forma supervisada como no supervisada. En líneas generales los estudios desarrollados pueden clasificarse en tres grandes grupos, en función del tipo de procedimiento de medida que utilizan [92]: procedimientos psicológicos, fisiológicos o motores-conductuales. El primer grupo de procedimientos, los de tipo **psicológico**, se basan en el análisis de auto-informes, verbales o escritos, a partir de los cuáles se evalúa el estado emocional del sujeto. Por su parte, los procedimientos de carácter **fisiológico** emplean sensores para poder capturar y medir las reacciones físicas ante un estímulo externo. Y por último, los procedimientos **conductuales** miden las expresiones y cambios en la conducta que pueden estar asociados a algún tipo de reacción emocional. Este último conjunto de procedimientos resulta ser menos invasivo que los dos anteriores, ya que no condiciona la participación del individuo en el momento de realizar la medida de las emociones.

En cuanto al objeto del análisis mediante procedimientos conductuales, se pueden destacar tres principales líneas de investigación: la detección de emociones en las expresiones faciales o corporales, en la expresión hablada y en la expresión escrita. En todos estos casos, la detección de emociones se realiza de manera indirecta; esto es, en condiciones de normalidad para el sujeto. Por ejemplo, a diferencia de los auto-informes sobre sus propias emociones que se utilizarían en un análisis de tipo psicológico, ahora se tratará de analizar textos escritos por el sujeto con otros fines, en los que se buscarán trazas emocionales que pudiesen estar presentes en dichos textos.

En el resto de esta sección se presentan diferentes técnicas y escenarios donde son utilizados los tres procedimientos de detección de emociones no invasivos ya mencionados (en texto, en habla y en expresiones faciales o corporales), así como algunas aplicaciones de detección multimodal. Asimismo, al final de cada apartado se indicarán brevemente algunos trabajos en torno a la clasificación emocional no supervisada en cada área.

### *2.1 Detección de emociones en textos*

En los últimos tiempos se han dedicado muchos esfuerzos en este área, debido en gran medida a la utilización masiva de medios de comunicación electrónicos. Las investigaciones y desarrollos se han dirigido fuertemente sobre la información producida desde internet, principalmente sobre las redes sociales, periódicos online, documentos digitalizados, o cualquier página Web con contenido textual que no se haya producido automáticamente. A continuación se presentarán diferentes técnicas para la detección de emociones en texto y casos de aplicación sobre los escenarios ya nombrados.

Un ejemplo de aplicación en esta línea es el de la identificación de la carga emocional en textos extraídos en titulares de periódicos online [75], utilizando la extensión afectiva del recurso léxico WordNet Affect [76]. Por su parte, Voll y Taboada [81] muestran la importancia de estudiar el impacto emocional contenido en los adjetivos contenidos en un texto, utilizando, en este caso, la herramienta SO-CAL (Semantic Orientation CALculator). Este procedimiento fue aplicado para estudiar el contenido emocional de los comentarios realizados por los usuarios de la web Epinion [1], encargada de recopilar opiniones de usuarios sobre su experiencia con determinadas marcas o establecimientos.

Con herramientas como Chat-See [18] puede verse la aplicación del análisis emocional evaluando el comportamiento emocional de las conversaciones entre estudiantes dentro de una plataforma de chats privados, realizando el etiquetado emocional de las palabras utilizadas dentro de las conversaciones. Continuando con el análisis en el entorno de la enseñanza, SentBuk [62] es una herramienta diseñada para recopilar y analizar emocionalmente los estados del Facebook publicados por los estudiantes, construyendo su propio léxico sentimental basado en la polaridad positiva y negativa con el fin de enriquecer el modelo de usuario de cualquier aplicación de e-learning. Así mismo, en [71] se crea una base de datos afectiva basada en word spotting, con el objetivo de analizar automáticamente los trabajos de los estudiantes a lo largo de un semestre o de uno o varios cursos.

La utilización de Facebook como medio para capturar información personal también es el objeto del trabajo de Adam Kramer [54] que, a lo largo de un año, retrata a gran escala la felicidad nacional de los Estados Unidos, evaluando las actualizaciones de estados de los perfiles de Facebook, asignando un peso emocional a dichos estados, y midiendo el porcentaje de positividad y negatividad.

También, en el ámbito de las redes sociales, Yassine y Hajj [89] proponen realizar un análisis emocional para determinar si lo escrito por el usuario contiene o no emoción, trabajando con un léxico de acrónimos populares utilizados en las redes sociales y en chats online. Continuando en este ámbito, el servicio de red social de Twitter es utilizado con mucha frecuencia por su particular característica de microblogging y uso de etiquetas (hashtags). Trabajando sobre Twitter, Bliss et al [13] presentan un medidor que evalúa la felicidad, llamado *Hedonometer* utilizado sobre Twitter para estudiar y evaluar cuáles son los temas que presentan un mayor grado de “felicidad” [36].

Asur y Huberman [7] presentan otro ejemplo de aplicación sobre Twitter, cuyo objetivo es predecir el comportamiento de la taquilla del cine, en EEUU, mediante la recopilación de *tweets* que contengan la etiqueta “movie”. En dicho trabajo, se correlaciona el nivel de felicidad con la aceptación de la taquilla en la semana que se estrena una película, realizando la medida de la



aceptación dentro del texto mediante la construcción de un clasificador sentimental basado en el paquete de análisis lingüístico LingPipe [6].

Por su parte, más que una propuesta concreta de análisis de reconocimiento o escenario de aplicación específico, Osherenko [64] propone que el análisis emocional se realice dividiendo los textos en párrafos y oraciones para facilitar su análisis, para esto, propone trabajar con tres diccionarios afectivos: Levin verbs [58], GI [74] y WordNet-Affect [76].

En otro entorno, y con el objetivo de publicar información emocional, Kamvar y Harris crearon We Feel Fine [48], un motor de búsqueda emocional que rastrea blogs, microblogs y redes sociales que contengan la frase “I feel” o “I am feeling” (“Me siento” o “Me estoy sintiendo”, respectivamente) con el objetivo de obtener el estado emocional de grupos de población según determinados parámetros, como pueden ser su ubicación, localidad, género, clima o edad.

Evidentemente, en el área de detección de emociones en textos existe un gran interés en crear plataformas y protocolos de trabajo que permitan desarrollos homogéneos; entre ellos están los diccionarios afectivos ANEW “Affective Norms for English Words” [14] y DAL “Dictionary of Affect in Language” [82]. Estos diccionarios, desarrollados originalmente para el idioma inglés, han sido posteriormente adaptados al español [69] [70] siguiendo, en ambos casos, el mismo procedimiento de traducción de las palabras originales y nueva ponderación.

La diferencia entre ambos yace en cuáles son las palabras con las que trabajan y en el tipo de parámetros de que utilizan para medir los niveles emocionales: ANEW contiene **1034** palabras y DAL **8742**. Únicamente **633** palabras están contenidas en ambos diccionarios. Con respecto a las medidas, ANEW utiliza la medida de los niveles de *Valence* (*Valencia* de agradable a desagradable) de *Arousal* (*Excitación* de calmado a tenso) y *Dominance* (*Dominio* de controlado a fuera de control) y DAL utiliza los niveles de *Pleasantness* (*Agrado* de agradable a desagradable), de *Activation* (*Activación* de pasivo a activo) y de *Imaginery* (*Imaginario* de fácil de imaginar a difícil de imaginar).

Otra diferencia notable entre ambos diccionarios afectivos es que ANEW ha sido utilizado como marco de trabajo para la medida de emociones sobre textos, a diferencia del contenido de DAL, que ha sido primordialmente utilizado para soportar la medida de gestos faciales y habla. Un estudio a fondo acerca de la comparación de estos diccionarios fue realizado por Cinthya Whissell [83].

Como ejemplo de escenarios de aplicación utilizando ANEW como base de medida afectiva podemos mencionar el trabajo de Mihalcea y Liu [63], dónde se utiliza ANEW para medir la felicidad a partir de la información contenida en blogs. Por su parte, Dodds y Danforth [35] también utilizan ANEW para medir los niveles de alegría y descontento contenidos en letras de canciones y en discursos anuales de los Presidentes de Estados Unidos. Klouman et al [51]

utilizaron ANEW para medir el positivismo del idioma inglés sobre cuatro diferentes corpus: Twitter, Google Books Project, The New York Times y letras de canciones.

Asociada al análisis de emociones, la clasificación emocional no supervisada o automática tiene gran importancia de manera natural. . En este sentido, Bhowmick et al [12] proponen un sistema de clasificación automático de noticias publicadas en periódicos dependiendo de su nivel emocional. Dentro de este mismo contexto, Kozareva et al [53] proponen un mecanismo para clasificar emocionalmente titulares de noticias online de forma automática. Siguiendo en el marco periodístico, pero desde otro punto de vista, Lin et al [59] y Bhowmick [11] proponen sendos procedimientos para realizar la clasificación emocional automática de noticias en base a los comentarios realizados por los lectores.

## ***2.2 Detección de emociones en el habla***

El estudio sobre la detección de emociones en el habla está presente en la literatura hace ya más de dos décadas. Entre las discusiones y análisis que se han realizado acerca de este tema, podemos destacar la revisión realizada por Schuller et al [72] o la presentada por Hirst [45]. En un contexto más amplio, es de destacar la labor de la asociación HUMAINE [2], que tiene como propósito estudiar las bases y la aplicación de la computación afectiva, incluyendo entre éstas el análisis emocional del habla y del discurso.

La existencia de herramientas comerciales capaces de reconocer las emociones del habla refleja el grado de madurez en este tipo de tecnologías. Un ejemplo de estas herramientas es ExSense [4], que ha sido definido como el “Analizador Vocal de Emociones”. El objetivo de ExSense es detectar el grado de stress en la voz del locutor, e indica que es capaz de determinar detectar si el locutor está mintiendo, confuso, concentrado e inclusive si está enamorado. También, una propuesta muy aplicable y útil en el ámbito comercial fue la realizada por Devillers y Vidrascu [34], que proponen etiquetar verbalmente las conversaciones para poder detectar las emociones en los diálogos telefónicos en call centers o asistentes telefónicos. En la misma línea, Lee y Narayanan [55] presentan otra propuesta de análisis emocional de la conversación aplicado sobre grabaciones realizadas en sistemas de call center, en la que detectan emociones negativas y no negativas en la conversación.

Entre las diferentes formas de realizar la medida de las emociones en el habla, podemos mencionar la propuesta por Yu et al. [91], cuyo objetivo fue crear un corpus emocional mediante la captura de pronunciaciones cortas en películas y adaptaciones teatrales de la televisión china. Esta propuesta utiliza técnicas de entrenamiento de algoritmos que puedan ser capaces de reconocer pronunciaciones que se encuentren dentro del conjunto de las emociones básicas.

Dentro de la detección emocional del habla, también se debe incluir la detección emocional de los registros vocales y, por tanto, el análisis emocional de la música, de la música cantada. En este sentido, se han propuesto diferentes enfoques y métodos para poder reconocer las emociones transmitidas dentro de las canciones o fragmentos de ellas [88] [85] [87] o [21]. En estos casos, las emociones se definen en términos de excitación (Excitado o calmado) y valencia (Positivo o negativo).

En otro orden de cosas, desde el punto de vista de las posibilidades de clasificación no supervisada, varios trabajos proponen la clasificación emocional del discurso y la voz para proveer un marco de trabajo en la medida y detección de emociones por medio del habla [80] [10] [86] [79]. Lee et al [57] proponen la clasificación de los estados emocionales basado en el modelado de fonemas; otros ejemplos de clasificadores emocionales del habla se pueden encontrar en los trabajos de Batliner y Huber [9] y Grimm et al [42].

### ***2.3 Detección de emociones en expresiones faciales y corporales***

A continuación se presenta una breve reseña de técnicas y aplicaciones dirigidas a la detección y análisis de las emociones transmitidas a través de los gestos faciales y corporales. Existen trabajos muy interesantes dentro de este área, como es el proyecto realizado por Castellano et al [24], cuyo objetivo es generar contenido audiovisual a través del movimiento y de los gestos faciales; es decir, asociar determinados colores y músicas a los movimiento del cuerpo y la cara, basándose en la información emocional transmitida por los gestos.

Desde el nivel experimental, existen estudios muy interesantes y prometedores que toman el estudio de la detección de emociones sobre las expresiones faciales desde diferentes puntos de vista; por ejemplo, en [77] se presenta un enfoque novedoso para entrenar un modelo probabilístico [44], con capas compuestas por diferentes corpus de expresiones faciales, teniendo como objetivo la generación de sistemas que sean capaces de generar estas expresiones faciales.

Azcarate et al [8] presentan un sistema para el reconocimiento automático de emociones a través de las expresiones faciales desplegadas en secuencias de video, además de estudiar la clasificación automática para el reconocimiento de dichas expresiones faciales. Las emociones básicas que se consideran en este trabajo son las sugeridas por Ekman et al [38]. Otro estudio muy interesante es el realizado por Friberg [41], que analiza las expresiones faciales que se presentan al tener contacto (sea de forma pasiva o activa) con la música. En dicho estudio, se presenta un analizador de expresiones emocionales en el marco de las interpretaciones musicales y movimientos corporales, cuyo objetivo era detectar la felicidad, la tristeza y la furia.

Por otra parte, Kim y André [50] sugieren el método de inducción musical para capturar las señales fisiológicas (a través de biosensores en el tórax, manos, abdomen y cuello) mientras el

sujeto escucha diferentes piezas musicales. También podemos identificar otro grupo de trabajos que estudian las emociones transmitidas (a través de los gestos faciales) por músicos cuando están interpretando el instrumento musical o al cantar, además del cambio físico que surge al momento de interpretar el solo [24] [32] [31] [33] [23].

Con respecto a la clasificación automática de gestos faciales, trabajos como los de Lyons et al [60], Padgett et al [65] o Hu et al [46] proponen diferentes enfoques para realizar la captura y clasificación de esta información.

#### ***2.4 Detección combinada de emociones (multimodal)***

La detección de emociones e implementación de sistemas afectivos combinados o multimodales también tiene una amplia cabida en este ámbito. Una de las áreas en que se puede observar es la educativa, como el caso del tutor inteligente o agente afectivo “Alice” [61]. Este agente forma parte de un sistema e-learning que tiene como propósito evaluar el estado afectivo de los estudiantes analizando sus expresiones faciales, su voz y sus textos. Por su parte, Conati y Maclaren [28] sugieren un modelo de usuario afectivo y probabilístico para permitir que un agente inteligente pueda reconocer múltiples emociones en los usuarios durante la interacción de los mismos en un juego educacional de ordenador.

Como referencia en estudios de investigación para la detección de emociones combinadas, se puede nombrar el proyecto europeo más importante dentro de la computación afectiva multimodal: HUMAINE. Su base de datos [37] ha sido utilizada en una gran cantidad de trabajos y fue creada con el propósito de poder almacenar información “emocional” en diferentes modalidades (texto, habla, gestos, etc.), además de almacenar técnicas de etiquetado para cada una de las modalidades. Además de la base de datos ya nombrada, también se utilizan otras herramientas dentro del proyecto, como el diccionario afectivo DAL [82].

También, en trabajos como los planteados por Calvo y D’Mello [19] o Cowie et al [29] se pueden encontrar análisis profundos acerca de modelos, métodos, técnicas y aplicaciones utilizadas para mejorar la experiencia persona-ordenador dentro del contexto del reconocimiento de emociones. De una forma más genérica, [56] y [20] sugieren una técnica dirigida a mejorar el reconocimiento de cualquier tipo de emociones.

Respecto a la combinación de técnicas concretas, Chuang y Wu [25] sugieren procedimientos para la combinación de la detección emocional del habla y del texto. Por su parte, el detector afectivo descrito por D’Mello y Graesser [30] combina los rasgos en la conversación con lenguaje corporal y características faciales, mientras el enfoque Conati [27] realiza una combinación de información explícita de la reacción emocional y efecto de dichas reacciones (detectada vía sensores filológicos).

Más allá de las técnicas utilizadas, la finalidad de la mayoría de los sistemas previamente mencionados es crear un modelo de usuario emocional – afectivo con el fin de complementar o mejorar sistemas previamente implementados.



### 3 Proceso de construcción de los conjuntos emocionales

---

En este apartado se presentan, en primer lugar, los conjuntos emocionales que se han utilizado como base para las clasificaciones que se muestran en el siguiente apartado. Estos conjuntos emocionales se han construido ad-hoc con motivo de esta investigación, y para ello nos hemos centrado en las cuatro emociones básicas definidas como tales a partir de la propuesta de Ekman[38], y soportadas por varios autores [92] [8]; esto es: *joy*, *sadness*, *anger* y *fear*. Luego, se realiza una comparación entre nuestros conjuntos de palabras emocionales y la versión de ANEW en español [69], Spanish-ANEW o S-ANEW de aquí en adelante, comparando nuestros resultados con los datos ofrecidos por S-ANEW.

Evidentemente, la decisión de construir un nuevo diccionario afectivo frente a S-ANEW fue una de nuestras principales decisiones de diseño, dado que S-ANEW es extensamente utilizado en otras aplicaciones. En este sentido, es de destacar que uno de nuestros principales objetivos era poder basarnos en un procedimiento extensible, que fuese fácil de mantener sin la utilización de grandes y costosos recursos. Con estas premisas claras procedimos a evaluar S-ANEW.

Cuando se iniciaron las investigaciones sobre el contenido de S-ANEW, se pudo notar que el proceso de adaptación del inglés al español consistía en la traducción literal al español de las 1034 palabras del original, dejando de lado sin evaluar muchas palabras importantes de uso frecuente en el idioma español y tomando en cuenta palabras que no parecen significar valores emocionales especialmente significativos, tales como *botella* (bottle), *rana* (frog), *hervir* (scalding), *inquieto* (troubled), *arrepentido* (repentant), *tenue* (subdued), *duro* (hard), *pizza*, *taxi*, etc. Además, la evaluación emocional de las palabras contenidas en S-ANEW implicó la participación de 720 personas para calificar los niveles de valencia, excitación y dominio de cada palabra. Si se plantease la extensión de S-ANEW, sería necesario repetir la calificación, lo que agregaría un coste importante a la evaluación del diccionario, sumándose a la dificultad añadida de realizar la propia extensión de los datos contenidos en el diccionario.

La diferencia entre nuestra propuesta y S-ANEW yace en que ellos trabajan con palabras que se utilizan en la vida cotidiana que de alguna forma despiertan emociones y han sido calificadas con el criterio subjetivo de un grupo de personas. En cambio nosotros proponemos trabajar con palabras que, a través de su significado (de diccionario), evocan una emoción sin la intervención subjetiva del criterio humano. Es decir que, al analizar emocionalmente cualquier tipo de texto, analizamos el grado de influencia que las palabras “evocadoras” de emoción tienen sobre un texto, sin importar el contexto textual en que este esté escrito (juegos de palabras, negaciones).

A continuación se presenta el proceso de construcción del enfoque propuesto, así como su comparación con S-ANEW:

- En una primera parte se presenta el proceso que se siguió para la identificación de las palabras que evocan las emociones y llenado de los conjuntos emocionales, en el marco de las cuatro emociones básicas.
- En la segunda parte se realizará la validación de lo propuesto por medio de la evaluación de las palabras contenidas en los conjuntos emocionales, correlacionándolas con las palabras de ANEW y su ponderación dimensional.

### 3.1 Construyendo los conjuntos emocionales

El enfoque presentado en este trabajo está basado en la recuperación de información a través de un procedimiento de traducción de inglés a español (*cross-linguistic information retrieval*). El punto de partida fue tomar las cuatro emociones básicas antes mencionadas y, a continuación, proceder a llenar estos cuatro conjuntos con palabras que, de alguna forma u otra, evocasen a cada una de dichas emociones. Por ejemplo, al llenar el conjunto de *anger*, éste debería dar cabida a las palabras que evocan a dicha emoción y así, igualmente, con cada una de las emociones.

El procedimiento para llenar cada uno de los conjuntos requiere de varios pasos importantes, para los que partimos de nuestras cuatro emociones básicas escritas en inglés: *joy*, *sadness*, *anger* y *fear*. A partir de ellas, el primer paso fue buscar en inglés los sinónimos<sup>2</sup> de cada una de las emociones, luego los adjetivos y por último palabras sugeridas por el diccionario. Posteriormente se realiza la traducción contextual de cada una de las palabras al español además de agregar palabras sugeridas a través de la traducción.

#### 3.1.1 Procedimiento de identificación de palabras que denotan emoción

Cuando se construye este tipo de enfoque *translingüístico* se adaptan los recursos y las técnicas del análisis sentimental del idioma inicial al nuevo idioma. Este tipo de construcción ofrece resultados muy interesantes independientemente del idioma en que se trabaje, aunque es muy común tomar el inglés como idioma inicial. En Brooke et al [17] y de Shi et al [73] se pueden revisar algunos ejemplos de uso de estas técnicas.

En el enfoque que se propone en este trabajo se ha decidido trabajar con el inglés como idioma inicial y el español como idioma final, basándonos en el recurso lingüístico online **bab.la** [3]. Este repositorio ofrece sugerencias de posibles traducciones para cada palabra, así como una lista de sinónimos de la misma y, también, oraciones completas en las que se muestra el uso de la palabra en el contexto de la traducción. Además, **bab.la** indica cuál es el origen de cada

---

<sup>2</sup> Enero del 2012, <http://bab.la>, Sinónimos obtenidos de © Princeton University



traducción sugerida. En nuestro caso, la mayoría de las traducciones recopiladas para este trabajo, i.e. relacionadas a las emociones, provienen de las traducciones oficiales de Parlamento Europeo. Todas las traducciones utilizadas para este trabajo han sido realizadas a lo largo de los meses de enero y febrero del año 2012, desde entonces el sitio **bab.la** ha incluido nueva información, además de mejorar la precisión de sus traducciones. Para este trabajo se verán reflejadas las traducciones recopiladas durante el periodo de tiempo antes mencionado.

Para poder conseguir los conjuntos emocionales en español, a continuación se explicará el procedimiento para la construcción de los cuatro conjuntos, en primera instancia en inglés, hasta llegar la construcción de los cuatro conjuntos finales en español.

Como ya se ha mencionado, partimos de la denominación de las cuatro emociones en inglés, creando cuatro conjuntos iniciales, en los que se incluyen los nombres y adjetivos asociados a cada una de las emociones. Como resultado se obtienen las siguientes duplas:

Joy – Joyful, Anger – Angry, Sadness – Sad, Fear – Afraid

Estas ocho entradas nos llevan a las listas de segundo nivel, que creamos agregando a cada conjunto los sinónimos resultantes de la búsqueda de cada término de la dupla dentro de **bab.la**, así como las palabras relacionadas proporcionadas por el diccionario. En la Tabla 3.1 se incluyen estas listas de segundo nivel, que constituyen la lista de emociones “expandida”, la que será el origen de los cuatro conjuntos emocionales.

### 3.1.2 Construcción de los conjuntos emocionales en español

El siguiente paso que se realizó fue construir cuatro conjuntos intermedios con la traducción al español de las palabras de cada uno de los cuatro conjuntos anteriores, trabajando con la herramienta más utilizada de **bab.la**: el traductor inglés-español. Para ello, se revisaron las traducciones sugeridas correspondientes a las duplas antes mencionadas (emoción – adjetivo). Además, también se incluyeron los verbos de las frases que sugiere el traductor con el fin de ayudar al usuario a dar contexto a la traducción.

Los cuatro conjuntos intermedios en español pueden ser consultados en el Apéndice 7.2. Como aclaración vale la pena mencionar que las palabras recopiladas han sido agregadas a cada conjunto sin ningún tipo de modificación y no se ha tenido interés en los posibles cambios de significado que pueda tener la palabra debido al contexto de la frase.

Tanto en los conjuntos en inglés como en español, se puede observar que hay algunas palabras que pertenecen a más de un conjunto. Por ejemplo, en la tabla del Apéndice 7.1 se observa como la palabra "distressing" está incluida tanto en el conjunto de *sadness* como en el de *fear*. Una interpretación de este resultado se asocia con el hecho de que esta palabra tiene la habilidad

de evocar más de una emoción y, por tanto, la utilizaremos en ambos conjuntos emocionales sin afectar los resultados finales.

El último paso fue crear los cuatro conjuntos emocionales finales, procesando las palabras para que poder ser utilizadas computacionalmente en el reconocimiento de emociones en texto. Para ello, se tuvo en consideración las posibles derivaciones de cada una de ellas.

Conviene señalar que este enfoque trabaja tomando las cuatro emociones básicas y su asociación con las palabras correspondientes, teniendo una selección semántica de las mismas basadas en su significado de diccionario, sin otra evaluación humana de dicha selección. Así, la carga emocional de cada palabra se asocia a su significado como tal, y no a posibles cambios de polaridad cuando la palabra se encuentra dentro de un contexto específico [22].

### **3.2 Validación del enfoque propuesto**

Con el fin de validar el enfoque propuesto, se ha realizado la comparación de las palabras contenidas en los cuatro conjuntos emocionales con las palabras contenidas en la adaptación al español de ANEW: S-ANEW [69].

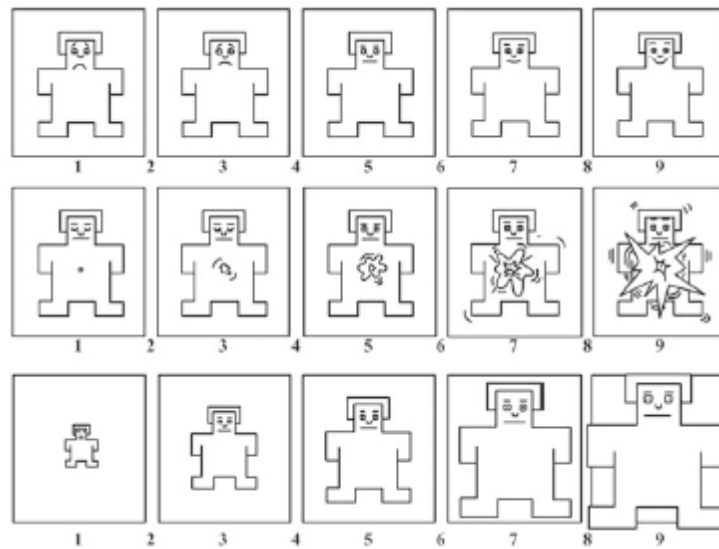
#### **3.2.1 ANEW y su adaptación al español**

El diccionario afectivo ANEW [14] fue desarrollado el año 1999 con el objetivo de proveer un conjunto de las 1034 palabras más utilizadas en la vida cotidiana que de alguna forma despiertan emociones en el mundo angloparlante. A cada una de estas palabras se le otorgó una calificación dimensional de las emociones, conceptualizando y organizando la emoción en tres dimensiones básicas subyacentes. Las dos dimensiones primarias utilizadas son: *valencia*, teniendo como rango desde placentero hasta desagradable, y *excitación*, teniendo como rango desde calmado hasta excitado. La tercera dimensión es *dominio* o *control* que oscila desde estar controlado hasta estar fuera de control. Esta última dimensión es considerada menos consistente que las dos dimensiones anteriores.

La calificación emocional de las 1034 palabras, con sus tres dimensiones correspondientes, fue realizada por 720 participantes, que utilizaron como parámetro de medida pictográfica no verbal el maniquí de autoevaluación SAM [16] llamado ScanSAM [15]. El resultado fue una base de datos normalizada que asocia los valores de valencia, excitación y dominio a cada una de las 1034 palabras en inglés, incluyendo su desviación estándar. Los valores de estas tres dimensiones oscilan entre 0 como valor mínimo y 10 como valor máximo.

En la adaptación española del ANEW, (S-ANEW), [69], la evaluación fue realizada con la colaboración de 720 participantes y con la versión adaptada para el ANEW del ScanSAM [14] para que pueda ser evaluado a través de lápiz y papel, con un ligero cambio en el rango de

calificación, siendo 1 el valor mínimo y 9 el valor máximo, tal y como se muestra en la Figura 3.1.



**Figura 3.1. SAM utilizado en S-ANEW. Escalas de autoevaluación para medir la: Valencia, Excitación y Dominio**

### 3.2.2 Comparando el enfoque propuesto en español con la adaptación española de ANEW

Con el fin de evaluar la adecuación de nuestra propuesta (los conjuntos emocionales), se realizó un estudio comparativo de ésta con S-ANEW, centrado, concretamente, en comparar como nuestra propuesta y S-ANEW clasifican las palabras de sus conjuntos y de qué tipo son las palabras contenidas en ambos.

La gran diferencia entre nuestra propuesta y S-ANEW yace en que ellos trabajan con palabras que se utilizan en la vida cotidiana que de alguna forma despiertan emociones y han sido calificadas con el criterio subjetivo de un grupo de personas. En cambio, en el caso de los conjuntos emocionales, la propuesta es trabajar con palabras que a través de su significado (de diccionario) evocan una emoción, sin la intervención subjetiva del criterio humano. Es decir que, al analizar emocionalmente cualquier tipo de texto, analizamos el grado de influencia que las palabras “evocadoras” de emoción tienen sobre un texto, sin importar el contexto en que esté escrito (juegos de palabras, negaciones, posible polisemia, etc.).

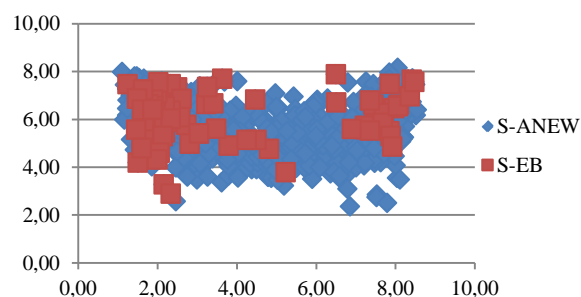
En otras palabras, los conjuntos emocionales propuestos en este trabajo tienen como propósito distribuir las palabras emocionales en los cuatro conjuntos que representan las cuatro emociones básicas, en cambio S-ANEW tiene como propósito asignar a cada palabra del conjunto tres valores: valencia, excitación y dominio, bajo la evaluación del criterio subjetivo de 720 participantes.

Ahora, la pregunta que se desea responder para poder llevar a cabo la comparación de estos dos enfoques es: *¿Todas (o la mayoría de) las palabras perteneciente a los cuatro conjuntos emocionales, comparten las mismas características que S-ANEW? O en otras palabras ¿Son ambos diccionarios consistentes uno con el otro?*

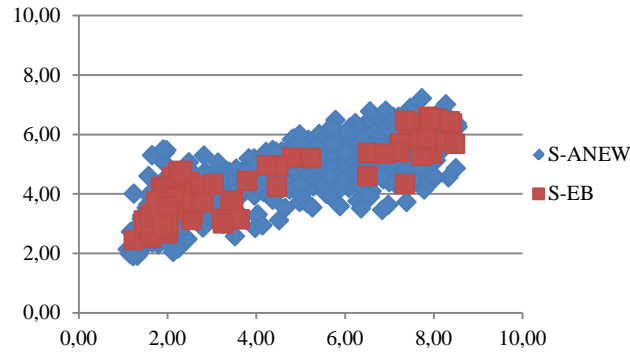
Para lograr este propósito se realizó la búsqueda de las palabras contenidas en ambos diccionarios, que resultarían ser la intersección entre ellos. Como resultado, se encontró que son 81 las palabras contenidas en ambas clasificaciones, perteneciendo 13 de ellas a dos conjuntos emocionales simultáneamente. Si se realiza la evaluación de la clasificación de S-ANEW con el enfoque propuesto, se puede decir que: 17 palabras evocaron *joy*, 19 evocaron *sadness*, 22 evocaron *anger* y 36 evocaron *fear*. El total suma 94 palabras, ya que hay 4 palabras que se incluyen tanto en *fear* como en *anger*, y otras 9 que se incluyen tanto en *fear* como en *sadness*.

En las tres Figuras que se muestran a continuación (Figuras 3.2, 3.3 y 3.4), los puntos azules representan todo el conjunto de S-ANEW, mientras que los puntos rojos representan las palabras de S-ANEW que están contenidas en los cuatro conjuntos emocionales en español. En la Figura 3.2 los ejes X vs. Y representan valencia vs excitación, respectivamente. En la Figura 3.3 se representa valencia vs dominio y, por último, en la Figura 3.4, se representa excitación vs dominio.

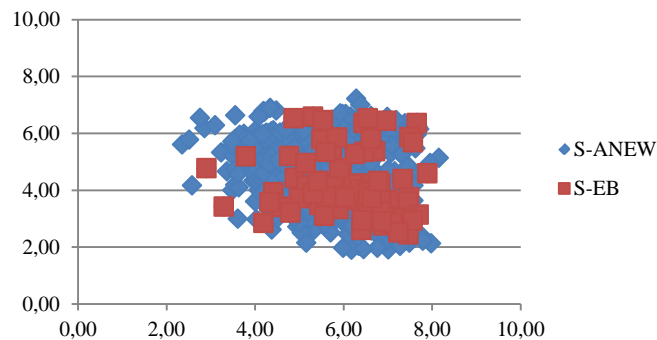
De las tres Figuras anteriores, los datos del conjunto emocional de la Figura 3.2 contienen los valores más significativos con respecto a patrones de distribución, esto se debe a que los datos extremos obtenidos en el eje de la valencia, a diferencia del eje de la excitación, son más homogéneos, con una tendencia general a tener valores altos. En segundo lugar podemos ver que los datos de la Figura 3.3, en cuyo eje de la valencia los datos se encuentran más esparcidos, pero manteniendo, aunque en bajo nivel, el patrón de distribución; en cambio, en el eje del dominio no se puede distinguir o extraer patrones acerca de la distribución de las palabras relacionadas. Finalmente la Figura 3.4, parece que no provee ninguna información válida acerca de la distribución de los valores de excitación y dominio con relación a las palabras del conjunto de emociones.



**Figura 3.2. Valencia vs. Excitación**



**Figura 3.3. Valencia vs. Dominio**



**Figura 3.4. Excitación vs. Dominio**

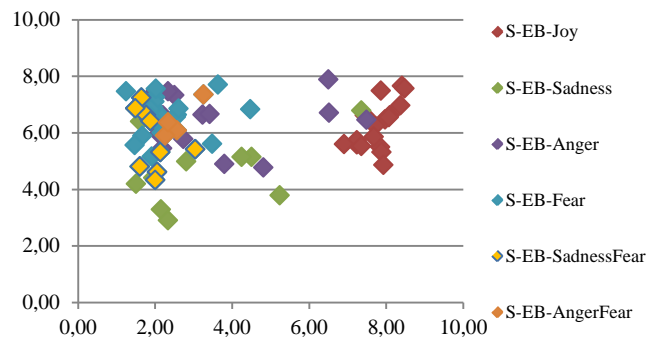
Estos resultados no son sorprendentes, dado que, por un lado la dimensión de Valencia (representada en las Figuras 3.2 y 3.3) es reconocida por ser la más significativa de las gráficas, siendo frecuentemente la única utilizada en las investigaciones relacionadas (y mencionadas anteriormente) de ANEW. Por otro lado, el significado de la dimensión de Dominio, no es mencionado comúnmente por su falta de consistencia dentro del estudio [69].

Por otra parte, en las Figuras 3.2 y 3.3 se pueden ver que las palabras del conjunto de emociones (puntos rojos) están concentradas en ambos extremos del gráfico. Por el mismo concepto de las tres medidas utilizadas por ANEW, los resultados anteriores no fueron sorprendentes. Por ejemplo, cuando se analiza la Figura 3.2 se pueden notar dos grandes secciones: por un lado las palabras que tienen valores bajos de Valencia o emotivamente “no agradables” tienden a tener valores altos de Excitación. Y por otro lado la mayoría de las palabras con altos valores de Valencia o emotivamente “agradables” también tienen altos valores de Excitación. Asimismo, la mayoría de las palabras con valor de Valencia intermedio o neutral, son también evaluadas con un valor intermedio o neutral de Excitación [69]. En consecuencia, era altamente probable que las palabras del conjunto de emociones se concentrasen en los extremos de la gráfica, como se puede apreciar. De hecho solo 5 de las 81 palabras contenidas en el conjunto de palabras tienen una Valencia evaluada entre 4 y 6.

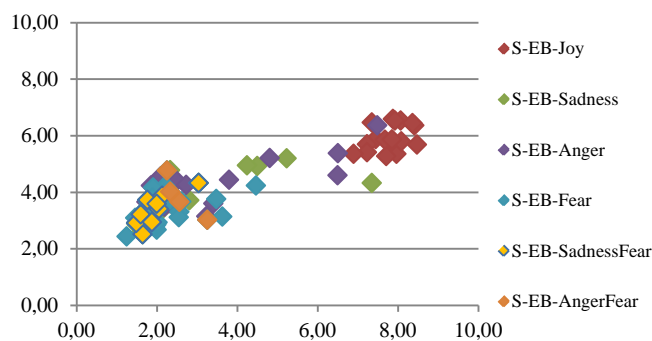
A continuación, las Figuras 3.5, 3.6 y 3.7 representan la misma relación que mostramos en las Figuras 3.2 a 3.4, pero en este caso las palabras de los conjuntos emocionales han sido coloreadas dependiendo al conjunto al que pertenecen. Las palabras repetidas, es decir, las palabras que pertenecen a más de un conjunto, son coloreadas y nombradas de forma distinta.

Las Figuras 3.5, 3.6 y 3.7 muestran relación entre alguna de las categorías de los conjuntos emocionales y las dimensiones de S-ANEW, especialmente las que representan los valores de Valencia, por ejemplo:

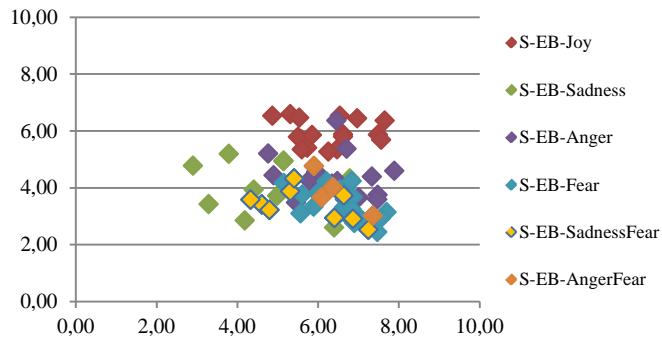
- Palabras del conjunto *joy*, todas tienen altos valores de Valencia.
- Palabras del conjunto *fear*, tiene una fuerte tendencia a tener valores bajos de Valencia, al igual que las palabras de los conjuntos mixtos *anger-fear* y *sadness-fear*.
- Aparentemente no se percibe algún patrón de distribución en los valores de Valencia en las palabras de *anger* y *sadness*.



**Figura 3.5. Valencia vs. Excitación**



**Figura 3.6. Valencia vs. Dominio**



**Figura 3.7. Excitación vs. Dominio**

Conclusiones similares se pueden extraer de los resultados relacionados con el valor de Dominio: existe una tendencia a que las palabras del conjunto *joy* tengan valores altos, mientras las palabras de los conjuntos *fear*, *anger-fear* y *sadness-fear* tienden a tener valores bajos.

### 3.3 Clasificación no supervisada de las palabras del conjunto de emociones a partir del contenido de S-ANEW

A partir de los resultados descritos anteriormente, el siguiente análisis tiene el propósito de comprobar si es posible construir clasificadores para un conjunto de palabras comunes, de ambas clasificaciones, utilizando técnicas de data mining. Explicado de otra manera: *¿Sería posible construir un clasificador que, dado un conjunto de valores pertenecientes a una palabra en S-ANEW (valencia, excitación y dominio), pudiera predecir a que categoría pertenece dicha palabra dentro de los conjuntos emocionales?*

Con la intención de contestar esta pregunta, se trabajó con la implementación J48 de árboles de decisión, proporcionada por la herramienta de aprendizaje automático Weka [84] y la exactitud de todos los clasificadores que se presentan en este trabajo se estimó utilizando *10-folds cross-validation* [52]. El dataset se construyó en base a las 81 palabras que tienen en común S-ANEW y el conjunto de emociones, tomando, a priori, las cuatro emociones básicas como las clases del clasificador. Sin embargo, pero como se mencionó anteriormente, existen palabras que pertenecen a más de un conjunto emocional, por este motivo se crearon dos nuevas clases llamadas: *sadness-fear* y *anger-fear*. Con la información necesaria se generó el árbol de clasificación, con un error estimado de 3,83%.

Dado que el árbol resultante es muy grande para ser mostrado en este documento, se mostrarán algunos ejemplos de las reglas que pueden ser derivadas del árbol.

$\text{Dominio} \leq 4,33 \ \& \ \text{Excitación} \leq 4,9 \rightarrow \text{Sadness}$

$\text{Dominio} > 5.1 \ \& \ \text{Valencia} \leq 7.44 \rightarrow \text{Joy}$

$\text{Dominio} > 5.1 \ \& \ 7.44 < \text{Valencia} \leq 7.49 \rightarrow \text{Anger}$

Dominio > 5.1 & Valencia > 7.49 → Joy

A la vista de estos resultados, es posible concluir que existe una clara correspondencia entre la descripción de las palabras propuestas por S-ANEW y la clasificación propuesta por los conjuntos emocionales. Existe además una interesante observación acerca del árbol de clasificación obtenido: la raíz del árbol ha resultado ser la dimensión de Dominio. Esto significa que el algoritmo de aprendizaje consideró que el valor de Dominio de una palabra es la característica que provee más información sobre el conjunto al que pertenece

Este resultado parece ser contradictorio a la percepción que se ha mostrado hasta ahora de la dimensión de Dominio, que aparentemente es menos significativa que las otras dos, Valencia y Excitación.

### **3.4 *Discusión***

Tal como se ha presentado en este trabajo, es posible concluir que existe una clara correspondencia entre la descripción de las palabras propuestas por S-ANEW y las palabras contenidas en nuestra propuesta. Este resultado nos motiva a examinar (en un futuro cercano) las palabras clasificadas en S-ANEW que no están contenidas en los conjuntos emocionales, para poder enriquecer el contenido de nuestros conjuntos. Por último, se puede concluir que nuestra propuesta ha demostrado que proporciona un mecanismo coherente capaz de construir conjuntos de palabras “evocadoras” de emociones asociadas a emociones concretas. Por otra parte, su enfoque *cross-lingüístico* permite que se aplique desde una perspectiva multilingüaje, con el fin de crear conjuntos emocionales en otros idiomas.



## 4 Análisis de la categorización automática de autores y obras literarias

---

En este apartado mostraremos cómo el análisis emocional de textos a partir de nuestro diccionario afectivo nos permite realizar la categorización emocional de trabajos literarios y conversaciones informales. Para ello se presentan dos estudios de diferente tipo:

En el primero, se intentará categorizar perfiles por medio del análisis emocional en tres escenarios: dos de ellos relacionados a la búsqueda de perfiles emocionales en la literatura y el tercero relacionado con la búsqueda de perfiles emocionales en conversaciones escritas.

Por su parte, el objetivo del segundo estudio es realizar la categorización emocional no supervisada, por medio del análisis emocional, de los poemas de Francisco Quevedo.

### 4.1 *Categorización emocional de perfiles emocionales de autores y conversaciones*

En esta sección se presenta la categorización de perfiles emocionales de tres escenarios concretos, tomando como emociones básicas *joy*, *sadness*, *anger* y *fear*:

- En el primer escenario se utilizará el análisis emocional de textos para identificar los diferentes perfiles que se pueden distinguir en los trabajos de cinco escritores de cuentos infantiles.
- El segundo escenario está enfocado en estudiar el resultado del análisis emocional de los trabajos de Gustavo Adolfo Bécquer (España, 1836 – 1870) y Mario Benedetti (Uruguay, 1920 – 2009). Ambos escritores fueron famosos por escribir tanto en verso como en prosa, y aquí se analiza parte de su producción en ambos géneros literarios
- Y por último, los mismos métodos aplicados en los dos anteriores escenarios serán aplicados a dos grupos de conversaciones escritas intercambiadas por medio de correos electrónicos.

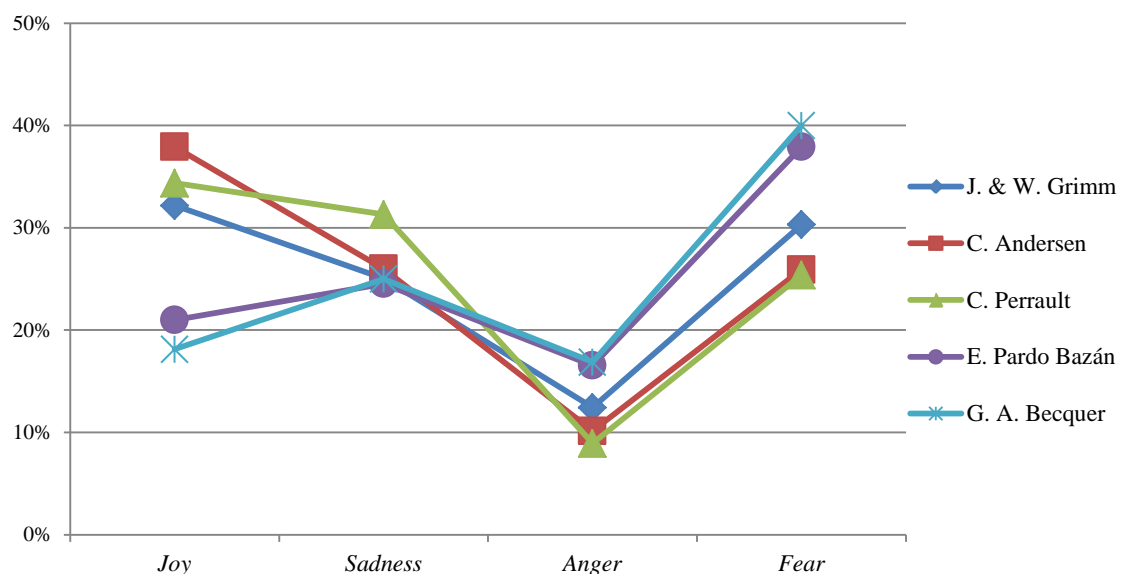
#### 4.1.1 Autores de cuentos infantiles

En este primer escenario se analizan emocionalmente cuentos infantiles de cinco escritores famosos por su gran trayectoria literaria; ellos son: Charles Perrault (Francia, 1628 –1703), Hans Christian Andersen (Dinamarca, 1805 – 1875), Gustavo Adolfo Bécquer (España, 1836 – 1870), Emilia Pardo Bazán (España, 1851 - 1921), Los hermanos Jacob Grimm (Alemania 1785-1863) y Wilhelm Grimm (Alemania 1786-1859).

De estos escritores se obtuvieron y analizaron emocionalmente entre 10 y 157 cuentos: Perrault 10, Bécquer 10, Pardo Bazán 16, los hermanos Grimm 119 y Andersen 157. Todos estos relatos

se obtuvieron de la biblioteca virtual “Ciudad Seva” [5]. La búsqueda y recopilación de los cuentos fue realizada entre los meses de enero y marzo del año 2012.

Los resultados de este análisis emocional lo podemos ver en la Figura 4.1, en la que se puede observar una clara división en dos estilos o clústeres emocionales. El primer clúster que puede verse en la figura es el que está formado por los tres escritores que concentran el porcentaje de *joy* más alto, siendo estos los **Hnos. Grimm, Andersen y Perrault**; el segundo clúster está formado por los dos escritores que concentran el porcentaje de *joy* más bajo, **Pardo Bazán y Bécquer**.



**Figura 4.1. Cuentos infantiles.**

Continuando con el análisis se puede observar que las dos emociones más poderosas entre estos cinco autores son *joy* y *fear*, marcando la tendencia emocional de los escritores. También se puede observar en la gráfica que el porcentaje de *joy* es inversamente proporcional al porcentaje de *fear*, concluyendo que mientras más alegres sean los cuentos menos miedo transmitirán. Si nos centramos en los valores de *anger*, ésta sería la tercera emoción más determinante, ya que, aunque obtiene los valores más bajos, también contribuye a modelar la "clusterización" de los resultados. Concretamente, podemos observar que los dos escritores que tienen valores más altos de *anger* son aquellos que tienen el *joy* más bajo, y viceversa. Mas no así en el caso de *sadness*, que cuatro de los cinco autores comparten prácticamente el mismo resultado.

Como resultado, el análisis realizado con nuestra propuesta de conjuntos emocionales nos ha dado la posibilidad de identificar marcados estilos emocionales en escritores del mismo género literario, otorgando así la posibilidad de realizar categorización de escritores a partir del análisis emocional de sus obras en un mismo género literario.

#### 4.1.2 G. A. Bécquer y M. Benedetti

En esta segunda parte se realiza el análisis emocional de obras escritas por G.A. Bécquer y M. Benedetti desde dos perspectivas: La primera consiste en analizar emocionalmente obras que correspondan al género literario de cuentos, que para este caso consiste en el análisis de 22 cuentos escritos por Bécquer y 20 cuentos escritos por Benedetti. La segunda consiste en analizar obras escritas en verso; en este caso, 15 rimas escritas por Bécquer y 20 poemas escritos por Benedetti. Las obras de Bécquer, así como las de Benedetti, elegidas para este trabajo están listadas en el Apéndice 7.3.

El propósito de este estudio era investigar si era posible identificar el tipo de tendencia emocional que manifiestan ambos autores y, también, comprobar si existía alguna influencia sobre el perfil emocional dependiendo del género literario en que estén escribiendo. Al igual que los cuentos, las obras de ambos escritores fueron obtenidas de la biblioteca virtual “Ciudad Seva” [5].

##### 4.1.2.1 Gustavo Adolfo Bécquer

En primer lugar se presenta el análisis de 22 cuentos escritos por Bécquer, de los cuales se obtendrán dos gráficas en las que se podrá observar cual es la tendencia emocional que Bécquer transmite al escribir sus cuentos. De la misma manera, se presentará el análisis de 15 rimas escritas por él mismo escritor, donde se realizarán los mismos procedimientos aplicados anteriormente a sus cuentos.

En la Figura 4.2 se presenta la proporción media de emociones para el total de los 22 cuentos de Bécquer analizados. Como se observa en la gráfica, en estos cuentos existe, en media, una fuerte tendencia hacia *fear*, emoción cuya presencia alcanza el 39.75% del total de emociones; otorgando a esta emoción el título de la emoción predominante. La segunda emoción predominante resulta ser *sadness*, la tercera *joy* y, por último, los valores más bajos han sido atribuidos a *anger*.

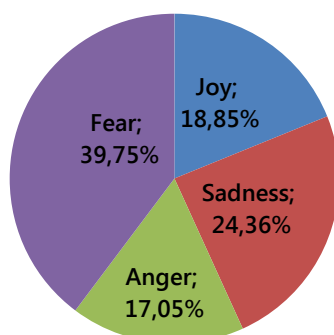


Figura 4.2. Proporción de emociones en 22 cuentos de Bécquer

Sin embargo, aunque la emoción predominante en los cuentos de Bécquer sea *fear*, existen algunos cuentos que difieren singularmente del patrón general, como es el caso del cuento

número 15, que tiene como emoción predominante a *joy* (los resultados de cada cuento se muestran en la Figura 4.3).

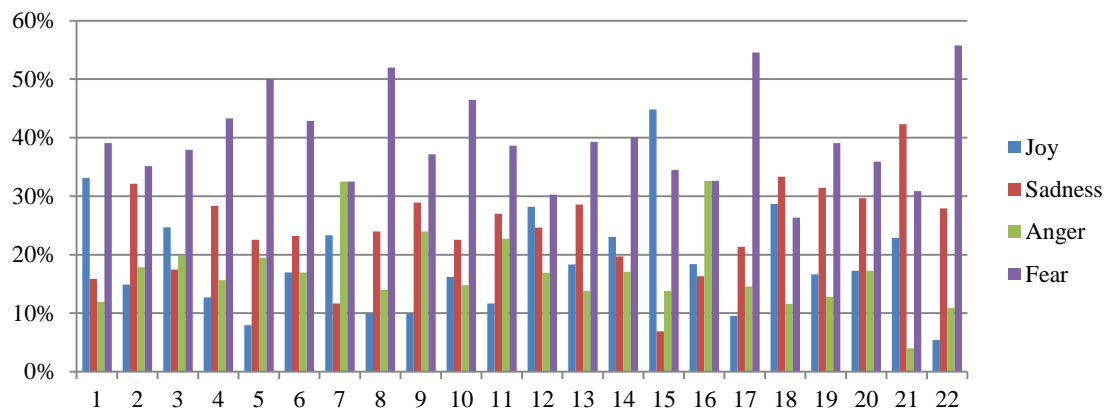


Figura 4.3. Cuentos de G. A. Bécquer

En la Figura 4.4 se muestra la agrupación de cuentos que tiene la misma proporción en la emoción. Para su elaboración, sobre el eje X se presenta la cantidad de cuentos y sobre el eje Y se presenta, en forma de barras, cada una de las emociones distribuidas en proporciones de 0% a 100%, en función del número de cuentos que contienen dicha proporción. En dicha gráfica se observa que *fear* es la emoción predominante, por cuanto ocupa posiciones más elevadas, mientras que *anger* siendo la emoción con la media más débil de las cuatro emociones, particularmente es la emoción que se encuentra presente en 18 de los 22 cuentos.

A continuación, en la Figura 4.5 se presentan los resultados de la proporción de emociones para las 15 rimas analizadas. A diferencia de sus cuentos de Bécquer, sus rimas muestran un comportamiento con mayor determinación emocional.

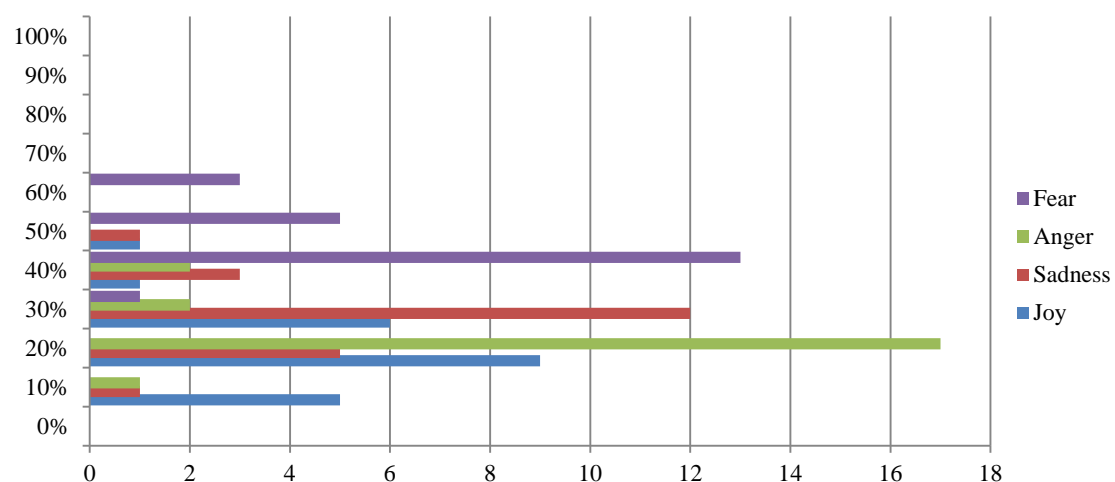
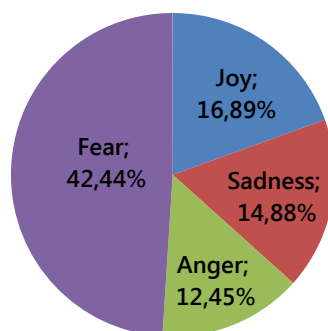


Figura 4.4. Agrupación de cuentos de Bécquer con la misma proporción emocional

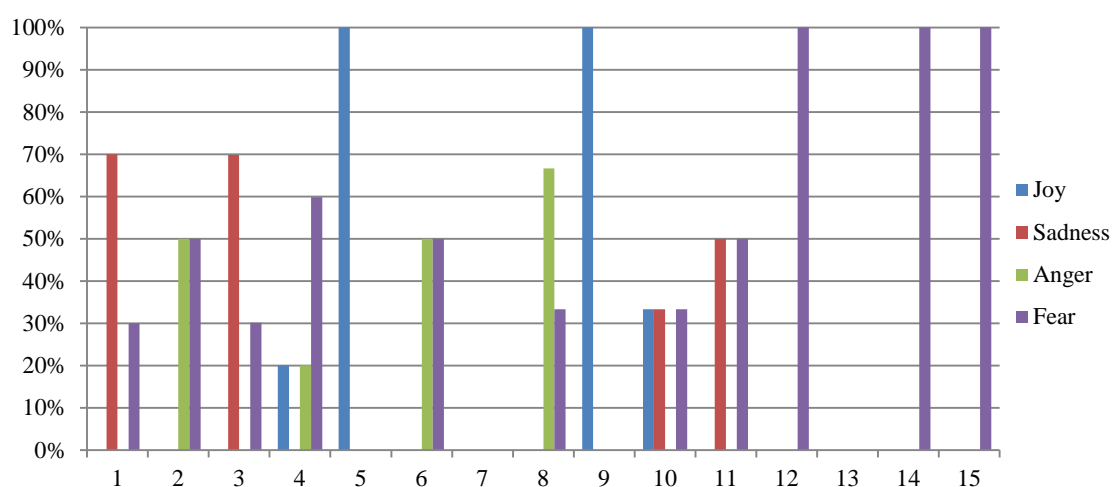
En esta ocasión se puede ver cómo, al igual que en los cuentos, las rimas de Bécquer tienen a *fear* como la emoción predominante y, también, se conserva el orden de proporción de las otras

tres emociones de forma descendente, *sadness*, *joy* y *anger*. Sin embargo, en este análisis *fear* alcanza casi la mitad del promedio total de las emociones, dejando valores muy bajos a las otras tres emociones.



**Figura 4.5. Proporción de emociones en 15 rimas de Bécquer**

Respecto a los resultados obtenidos por cada una de las rimas, en la Figura 4.6 pueden observar algunos casos particulares; como en el de las cinco rimas en las que solo se ha detectado una emoción: *joy* en las rimas 5 y 9, y *fear* en las rimas 12, 14 y 15. Además puede verse que existe la situación contraria en donde dos de las quince rimas carecen totalmente de emociones: las rimas 7 y 13.



**Figura 4.6. Rimass de G. A. Bécquer**

Al igual que en caso de los cuentos, en la Figura 4.7 se muestra la agrupación de rimass que tienen la misma proporción de emociones: en el eje X se representa la cantidad de rimass y, en el eje Y, en forma de barras, cada una de las emociones distribuidas en proporciones de 0% a 100%, como función del número de rimass que contienen dicha proporción. De la gráfica se observa que 10 de las 15 rimass carecen completamente de *anger*, *sadness* y *joy*, quitando gran relevancia a estas tres emociones y confirmando que la emoción predominante es claramente *fear*, dejando de lejos a *joy* en el segundo lugar.

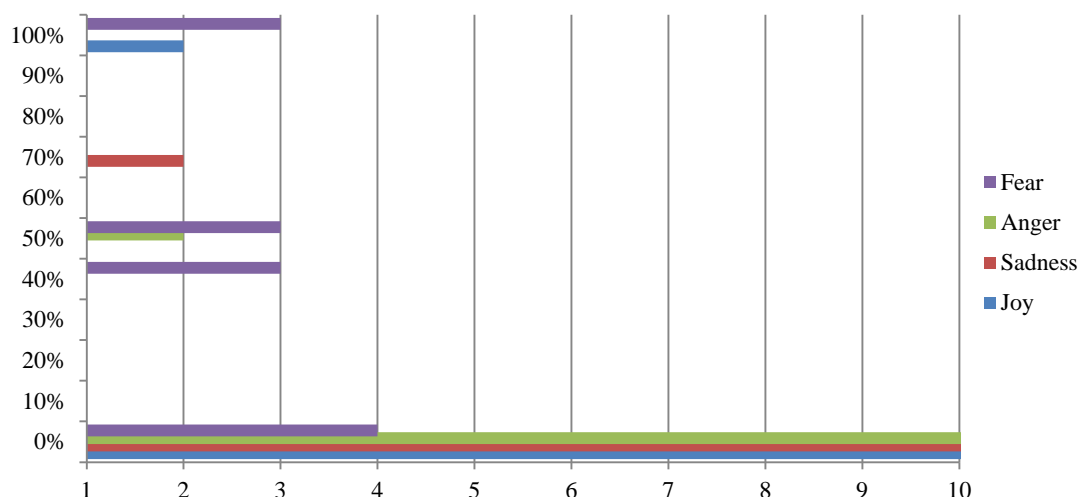


Figura 4.7. Agrupación de rimas de Bécquer con la misma proporción emocional

#### 4.1.2.2 Mario Benedetti

En este apartado se presenta el análisis emocional de 20 cuentos escritos por Benedetti, de los cuales se obtendrán dos gráficas, cuyo objetivo es poder distinguir cual es la tendencia emocional que transmite al escribir cuentos. Al igual se presentará el análisis de 20 poemas escritos por él mismo escritor, donde se realizarán los mismos procedimientos aplicados a sus cuentos.

En la Figura 4.8 se presentan los resultados de la proporción de emociones para los 20 cuentos analizados. Observando el comportamiento de las cuatro emociones en cada uno de los cuentos, se puede notar que la tendencia más fuerte es de *fear*, destacándose esta emoción como la predominante, con un cercano segundo lugar a *joy*, en tercer lugar a *sadness* y, por último, los valores más bajos han sido atribuidos a *anger*.

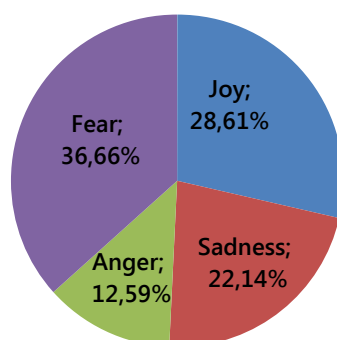


Figura 4.8. Proporción de emociones en 20 cuentos de Benedetti

Al igual que en caso de Bécquer, en la Figura 4.9, se puede advertir que aunque la emoción predominante sea *fear* existen cuentos puntuales que difieren del patrón general, como sucede con el cuento número 18, que tiene como emoción predominante a *joy* y, a la vez, es el cuento

que tiene el valor (de *joy*) más alto de los 20 analizados. Continuando con el análisis de los resultados del análisis de los cuentos de Benedetti, las emociones *sadness* y *anger*, además de ser las emociones más “débiles” de las cuatro, comparten en común el hecho de estar ausentes en algunos cuentos.

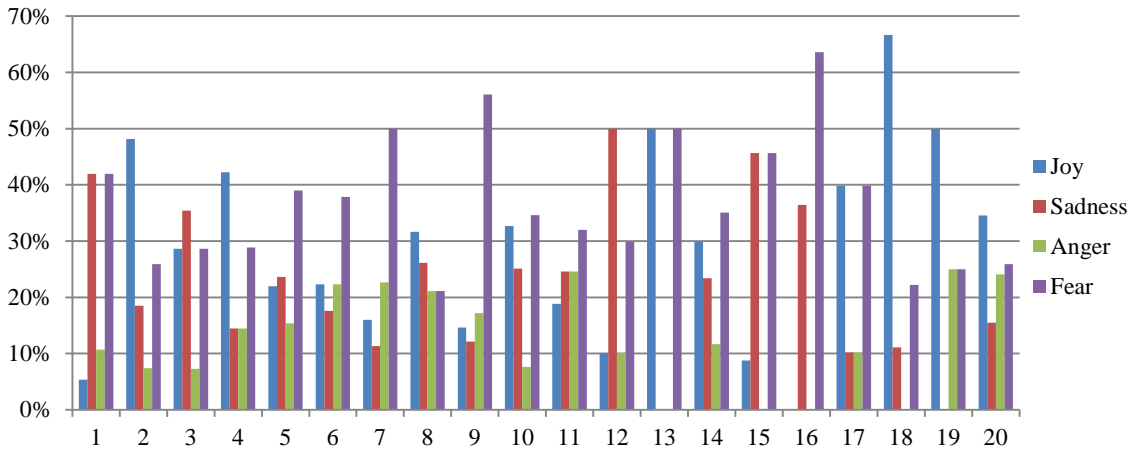


Figura 4.9. Cuentos de Mario Benedetti

Como en los casos anteriores, en la Figura 4.10 se muestra la agrupación de cuentos que tiene la misma proporción de cada emoción. En la gráfica se observa que *fear* es la emoción predominante por cuanto que ocupa posiciones más elevadas. Cabe observar que, aunque *sadness* y *anger* estén presentes en la mayor cantidad de cuentos, éstas se encuentran en las dos últimas posiciones. Eso quiere decir que están presentes en la mayoría de los cuentos pero es una presencia emocionalmente débil.

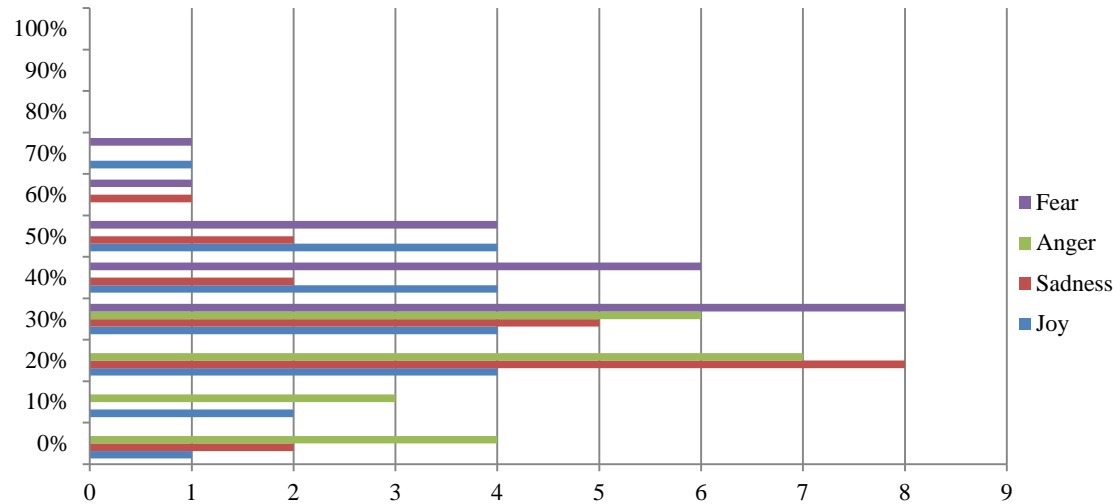
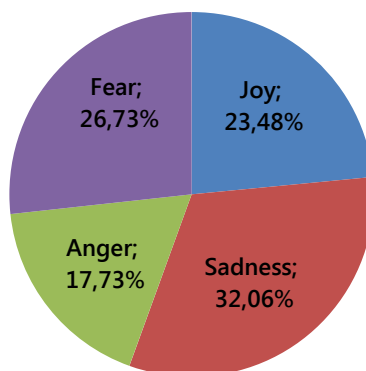


Figura 4.10. Agrupación de cuentos de Benedetti con la misma proporción emocional

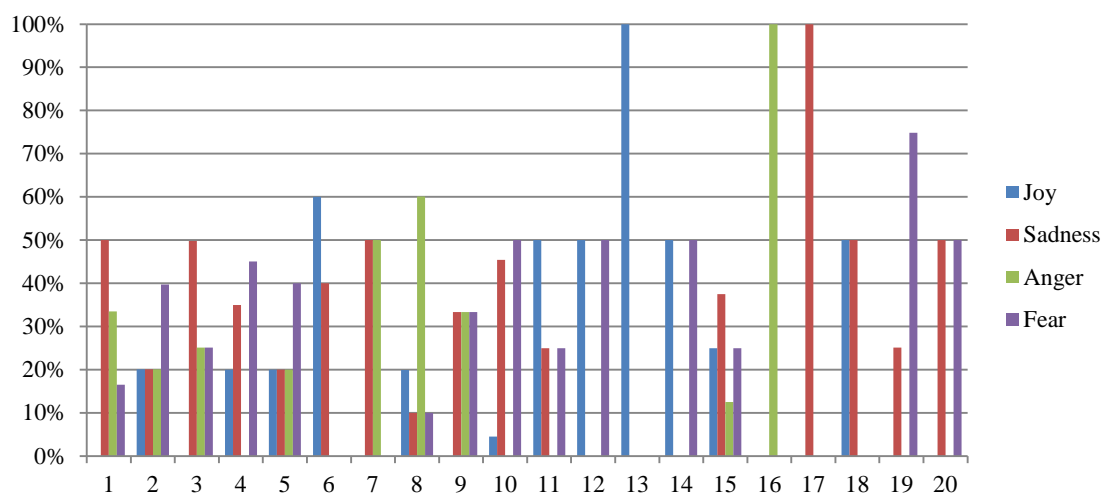
En cuanto a los poemas de Benedetti, en la Figura 4.11 se presentan los resultados medios de la proporción de emociones para los 20 poemas analizados. Observando el comportamiento de

Benedetti, se puede decir que a comparación de los resultados obtenidos en Bécquer y los cuentos de Benedetti, estos resultados son los que tienen mayor “equilibrio” emocional. Esto se puede ver en los resultados de *sadness*, que aunque en este caso sea la emoción predominante, no lo hace con una gran diferencia en comparación con los resultados de *fear* y de *joy*. Dejando en último lugar (como ha sucedido en todos los casos anteriores) los resultados de *anger*.



**Figura 4.11. Proporción de emociones en 20 poemas de Benedetti**

En la gráfica que se presenta en la Figura 4.12, los resultados para cada una de los poemas, se observa el caso particular de *anger* que aunque sea la emoción más débil de las cuatro en el promedio total, es la única existente un poema concreto, el número 16, que transmite el 100% de esta emoción, saliendo por completo del patrón emocional general.



**Figura 4.12. Poemas de Mario Benedetti**

Por último, en la Figura 4.13 se muestra, al igual que en los casos anteriores, la agrupación de poemas que tienen la misma proporción en cada emoción. En esta ocasión, se puede observar que *anger*, aunque esté presente en la mayoría de los cuentos analizados, el porcentaje de emoción que transmite es muy bajo.



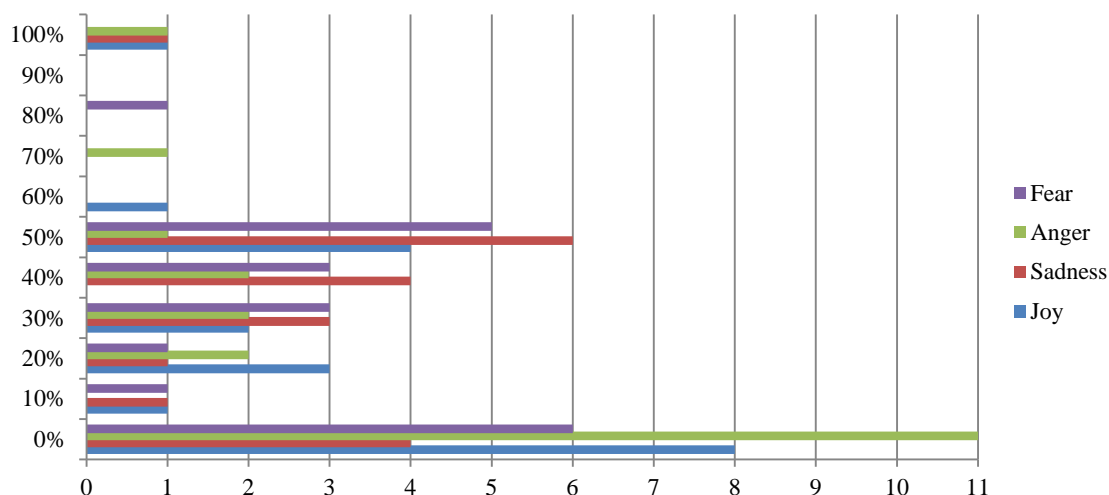


Figura 4.13. Agrupación de poemas de Benedetti con la misma proporción emocional

#### 4.1.2.3 Discusión de las obras de Bécquer y Benedetti

Comparando los niveles emocionales ya analizados con los valores máximos de cada emoción incluidos en la Tabla 4.1, se puede concluir que ambos escritores transmiten las emociones de forma más intensas cuando escriben en verso: las rimas en el caso de Bécquer y los poemas en el caso de Benedetti.

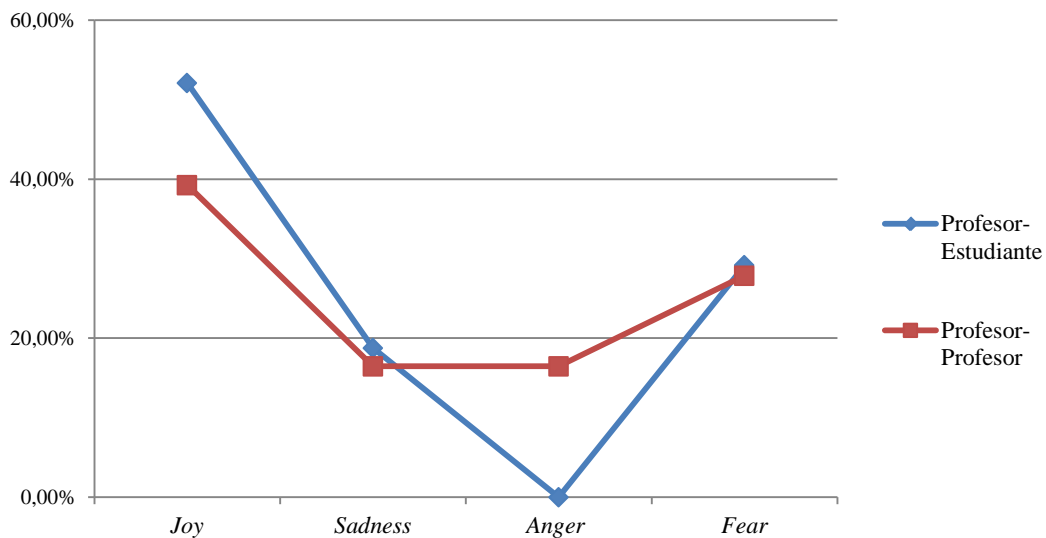
En conclusión, se puede decir que ambos escritores tienen mucha semejanza con respecto a la emoción predominante (*fear*), a excepción de los poemas de Benedetti, que tienen como emoción predominante a *sadness* pero con un margen muy pequeño con respecto a *fear*.

	<i>Joy</i>	<i>Sadness</i>	<i>Anger</i>	<i>Fear</i>
<b>Bécquer Cuentos</b>	44,83%	42,29%	32,63%	55,78%
<b>Bécquer Rimas</b>	100,00%	70,00%	66,67%	100,00%
<b>Benedetti Cuentos</b>	66,67%	50,00%	25,00%	63,57%
<b>Benedetti Poemas</b>	100,00%	100,00%	100,00%	74,87%

Tabla 4.1. Valores máximos de cada emoción

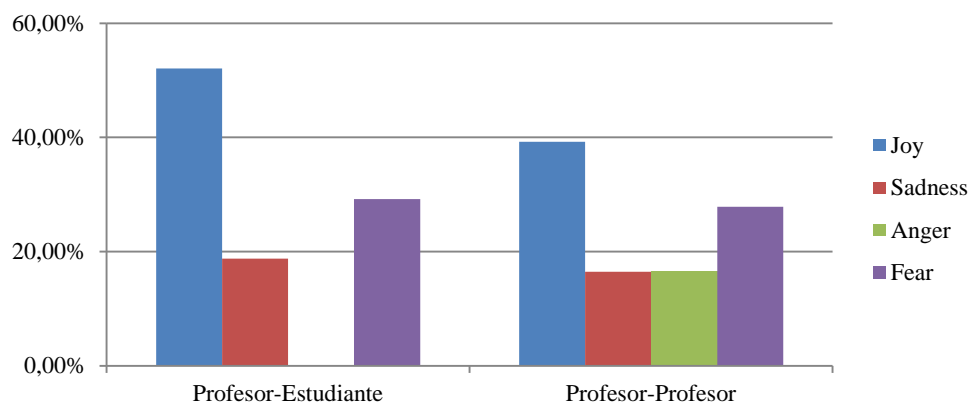
#### 4.1.3 Categorización de conversaciones en un contexto específico

El propósito de esta sección es mostrar el análisis de un escenario completamente diferente, analizando información contemporánea escrita de manera informal en un contexto específico. En este trabajo el contexto elegido ha sido el intercambio de correos electrónicos en un ambiente académico. Para lograr este propósito, se obtuvieron e-mails de dos tipos de conversaciones: el primero está compuesto por la compilación de 20 e-mails intercambiados entre dos profesores y el segundo contiene 20 e-mails intercambiados entre un profesor y sus estudiantes, en ambos casos, el intercambio de e-mails fue realizado en la misma asignatura y a lo largo del periodo académico.



**Figura 4.14. Porcentaje emocional del intercambio de e-mails**

Como puede verse en las Figuras 4.14 y 4.15, en los e-mails intercambiados entre profesor y estudiante, *joy* es la emoción que alcanza los niveles más altos, además puede notarse la completa ausencia de *anger* (que en español significa tanto *furia* como *pasión*) en las conversaciones; este hecho se puede atribuir al hecho de que son correos intercambiados con el único propósito de tocar temas que solo conciernen a la asignatura, es decir correos impersonales ausentes de “pasión” o “entusiasmo”.



**Figura 4.15. Análisis emocional del intercambio de e-mails**

Sin embargo, en el caso de las conversaciones entre ambos profesores los niveles de *joy* bajan pero *anger* aparece en escena, igualando su valor con *sadness* aunque se mantenga *joy* como la emoción con el nivel más alto, evidenciándose una mayor estabilidad emocional en las conversaciones realizadas entre compañeros de trabajo.

Como conclusión se puede decir que *joy* y *anger* son las emociones que realmente marcan la diferencia en ambas conversaciones.

## 4.2 Categorización emocional, no supervisada, de los poemas de Francisco Quevedo

Francisco Gómez de Quevedo Villegas y Santibáñez Cevallos, conocido por Francisco de Quevedo (1580-1645) fue uno de los escritores más prominentes del siglo XVII, la famosa *Edad de oro de la literatura española*. Aunque a lo largo de su vida escribió tanto novelas como poesía, su producción poética fue la más numerosa y la más reconocida, (escribiendo alrededor de 900 poemas). Sus poemas nunca fueron publicados oficialmente, pero fueron famosos por su distribución informal. A los pocos años de su fallecimiento algunos de sus poemas fueron editados en diferentes volúmenes por José Antonio González de Salas en el año 1648, luego en el año 1670 una nueva edición fue realizada por su sobrino Pedro Aldrete Quevedo y Villegas.

En los siglos siguientes se han realizado muchos estudios sobre los trabajos de Quevedo desde diferentes líneas de investigación, una de ellas relacionada con la compilación y edición de sus poemas y, en varios casos, han sido realizadas desde el punto de vista emocional de cada editor. Dos ejemplos muy conocidos son la edición de José Manuel Blecua Tejeiro (1913 – 2003) [66] y la edición bilingüe, español-inglés, realizada por Christopher Johnson [67].

Como se podría esperar, los poemas que se incluyen en cada edición, así como la clasificación (emocional) de los poemas, difiere entre cada editor. Por ejemplo, la edición de Johnson incluye solo 46 poemas bajo 8 categorías: Metafísicos, Heráclito cristiano, Poemas morales, Poemas líricos, Elogios y Epitafios, Amorosos, Canta sola a Lisi y Poemas satíricos y burlescos [67]. Por otra parte, en la edición de Blecua el número de poemas recolectados es de 185 bajo 4 categorías: Amorosos, Canta sola a Lisi, Satíricos y Poemas filosóficos, morales y religiosos [66]. A pesar de la diferencia entre el número total de poemas recolectados por cada editor, está claro que existe cierta correlación entre ambas clasificaciones. Concretamente son 3 las categorías en las que ambos coinciden: Amorosos, Canta sola a Lisi y Satíricos; los poemas de las 5 categorías restantes de la edición de Johnson pueden ser encontrados en la última clasificación de la edición de Blecua: Poemas filosóficos, morales y religiosos.

Viendo la similitud entre ambas clasificaciones, se ha decidido analizar emocionalmente los poemas escogidos en la edición de Blecua, principalmente por la cantidad de poemas clasificados (teniendo 185 vs. 46), pero también por las cuatro categorías propuestas. Estas categorías han sido tomadas como parámetro de medición emocional, a excepción de “Canta sola a Lisi” que, a priori, es la única difícil de clasificar. Esta última categoría, “Canta sola a Lisi”, contiene un conjunto de poemas dedicados a una mujer a la que se refiere el autor con el nombre ficticio de Lisi. Filólogos dedicados a estudiar los poemas de Quevedo no han sido capaces de responder al interrogante de si se trata de una mujer real o ideal o, incluso, si es una combinación de ambas. Dejando a un lado su existencia, la diferencia emocional con los poemas Amorosos parece ser un tema tan subjetivo como la existencia de la misma Lisi. Un ejemplo de

ello lo planteó Young [90], argumentando que los poemas dirigidos a Lisi difieren muy poco, por no decir nada, de los poemas Amorosos de Quevedo y por consecuencia él no hace ninguna distinción entre estas dos categorías.

Al margen de los problemas con dicha categoría, una vez decididos los poemas y sus respectivas clasificaciones, el siguiente paso consistió en analizar y obtener las cuatro emociones básicas de cada uno de los 185 poemas, que constituye nuestro conjunto de datos (dataset) inicial. Como se mencionó al inicio del documento, uno de los objetivos de este trabajo es probar si es posible construir un clasificador emocional automático capaz de imitar la clasificación de los poemas de Quevedo editada por Blecua. Además existe un particular interés en observar si los poemas de Canta sola a Lisi, son clasificados automáticamente o no como Amorosos.

#### 4.2.1 Categorización de los datos originales

En este trabajo nuestro interés es comprobar si se pueden utilizar técnicas de data mining para poder construir un clasificador que, dado un poema de Quevedo con sus cuatro emociones básicas analizadas, sea capaz de clasificar dicho poema, de forma automática, en una de las clasificaciones de Blecua. En otras palabras, se quiere comprobar si se puede realizar una clasificación emocional automática que sea consecuente con la categorización realizada por Blecua.

Bajo este contexto, se ha decidió abordar dos cuestiones: la primera consistiría en comprobar si las categorías originales de la clasificación de Blecua se podían distinguir a través del análisis emocional de sus poemas y, la segunda, en explorar diferentes técnicas de aprendizaje automático para encontrar el algoritmo que fuese capaz de producir el mejor resultado para el dataset dado.

Son varias las preguntas que surgieron al abordar la tarea anterior, la primera fue: *¿la clasificación de Blecua puede ser explicada en términos de patrones sentimentales?* Es decir *¿Los sentimientos detectados en cada poema podrían ser capaces de explicar la diferencia entre los poemas Amorosos y Canta sola a Lisi?* La segunda pregunta fue: *¿Y si se realiza la combinación de algunas de las categorías originales se podrían conseguir mejores resultados en términos del análisis sentimental?* O en otras palabras, *¿Se podría mejorar el clasificador eliminando una de las cuatro categorías?*

Con el fin de evitar diferencias producidas por los algoritmos de aprendizaje, en este primer estudio, solo se ha trabajado con Árboles de Decisiones [68]. Esta técnica es lo suficientemente simple para permitir pruebas eficientes sobre diferentes configuraciones; particularmente se ha trabajado sobre la implementación J48 de los Árboles de Decisiones provista por Weka [84].

### 4.2.2 Resultados

A continuación se describen las variaciones creadas a partir del dataset inicial para poder producir los resultados necesarios.

#### 4.2.2.1 Dataset original: 185 instancias y 4 clases

El dataset original está compuesto por 185 poemas divididos en 4 categorías diferentes: Amorosos, Canta sola a Lisi (Lisi ahora en adelante), Satíricos y Filosóficos-Morales-Religiosos (FMR ahora en adelante). Con este dataset, se construyó un árbol clasificador con una exactitud del 56,22%. La exactitud de todos los clasificadores utilizados en este trabajo fue estimada utilizando *10-fold cross validation* [52].

Aunque la exactitud nos haya dado un resultado muy cercano al 50% (56,22%), se puede ver que este tipo de clasificadores, basado en la detección de emociones, es más del doble de efectivo que el clasificador con una selección *uninformed* o que no está basado en ningún parámetro, que tendría un 25% de probabilidad para poder obtener la clase correcta de un poema.

En la siguiente Figura (Fig. 4.16) se puede ver el árbol de clasificación resultante construido con las 185 instancias (poemas) divididos en sus 4 clases correspondientes, donde cada poema posee los cuatro valores resultantes de las cuatro emociones básicas: *joy*, *sadness*, *anger*, *fear*.

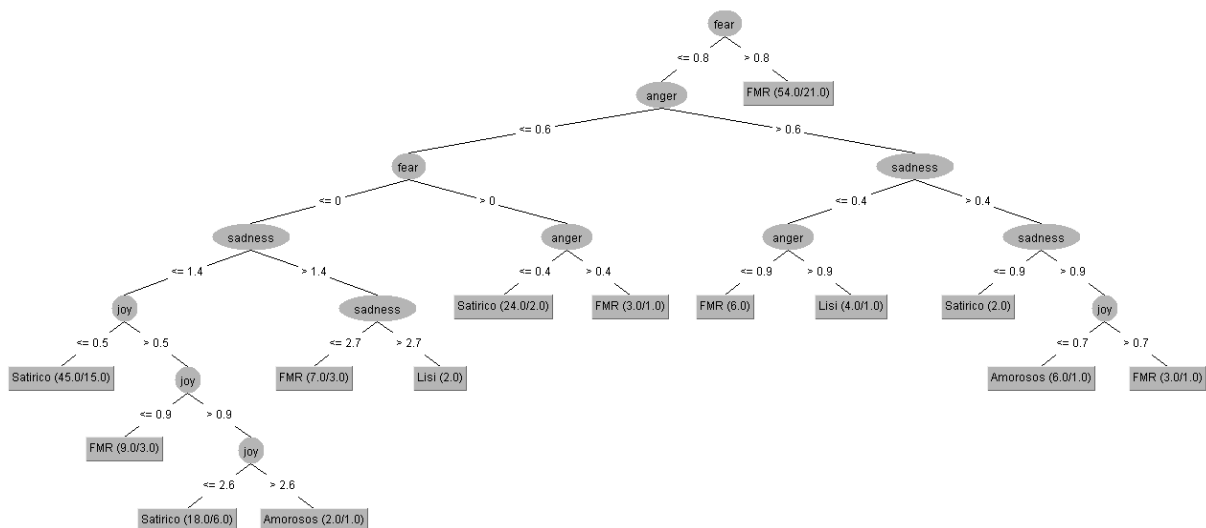


Figura 4.16. Árbol de clasificación con el dataset original: 185 instancias, 4 clases

La categorización que representa este árbol, aunque haya resultado ser un árbol relativamente pequeño (13 nodos internos y 14 hojas), parece refrendar, en el caso de los poemas de Lisi, lo que podría ser nuestra idea original; es decir, que son poemas con alto valor de *anger* (furia o pasión), bajo *fear* y bajo *sadness*.

Sin embargo, esta clasificación podría estar sesgada debido a que el número de instancias en cada categoría no estaba equilibrado.

La distribución de cada clase puede ser vista en la Tabla 4.2. En este tipo de situaciones, las clases con menos instancias tienden a ser ignoradas, por ejemplo, cuando se analiza la matriz de confusión en la Tabla 4.3 se puede ver que la mayoría de las instancias de Lisi y Amorosos han sido clasificados como FMR o Satíricos.

Clase	Valor
Lisi	22
Amorosos	21
Satírico	73
FMR	69

**Tabla 4.2. Distribución original de las clases**

El fin de la matriz de confusión es poder visualizar cuantas instancias de cada clase han sido clasificadas correctamente y cuáles de forma errónea. En este caso la cantidad de clasificaciones correctas pueden ser vistas en la diagonal de izquierda a derecha.

Clase Real	Clase Pronosticada			
	Lisi	Amorosos	Satírico	FMR
Lisi	<b>5</b>	0	6	11
Amorosos	2	<b>4</b>	7	8
Satírico	1	2	<b>58</b>	12
FMR	10	5	17	<b>37</b>

**Tabla 4.3. Matriz de confusión para el Clasificador 4.2.2.1**

También se pueden ver los valores-F en la Tabla 4.4:

Clase	Valor-F
Lisi	0,25
Amorosos	0,25
Satírico	0,72
FMR	0,54

**Tabla 4.4. Valores-F para el clasificador 4.2.2.1**

#### 4.2.2.2 Resampling: 740 instancias y 4 clases

Con el fin de evitar el sesgo por la distribución irregular de los poemas, se aplicó un filtro con un nuevo muestreo o Resampling [52]. Este filtro obtiene instancias al azar del dataset original y crea un nuevo dataset; si se especifica un sesgo, en lugar de elegir instancias aleatoriamente, a cada clase le será asignado un peso, en ese caso la distribución final del nuevo dataset podría ser diferente del original.

En este caso, el sesgo se configuró para que el nuevo dataset tendiera a poseer una distribución uniforme de las instancias de cada clase. Al mismo tiempo, con el fin de tener suficientes instancias para todas las clases, el filtro fue configurado para generar una muestra 400% más grande que la original. De otra manera para lograr el efecto principal del *resampling* se tendría que quitar instancias de las clases mayoritarias. La distribución del dataset resultante de aplicar el filtro anterior se puede ver en la Tabla 4.5.

Clase	Valor
Lisi	189
Amorosos	167
Satírico	169
FMR	215

Tabla 4.5. Distribución de las 4 Clases con Resampling

Con este nuevo dataset se construyó un nuevo árbol de clasificación con una exactitud del 75,13% (de la misma forma utilizando 10-fold cross validation). Este nuevo clasificador no solo mejora el rendimiento con respecto al original, sino que produce una clasificación más justa de las clases, como puede verse en la matriz de confusión de la Tabla 4.6.

Clase Real	Clase Pronosticada			
	Lisi	Amorosos	Satírico	FMR
Lisi	<b>179</b>	4	0	6
Amorosos	37	<b>117</b>	10	3
Satírico	38	5	<b>116</b>	10
FMR	46	14	11	<b>144</b>

Tabla 4.6. Matriz de confusión para el Clasificador 4.2.2.2

Además de proveer la información explicada anteriormente acerca de las matrices de confusión, en el caso de la Tabla 4.6 se puede ver que existe una pequeña confusión entre las clases de Lisi y Amorosos; a priori se ha pensado que, como los poemas de ambas clases están relacionadas con el amor, sería difícil de separar una de la otra, pero la matriz de confusión contradice lo que a priori se asumió.





#### 4.2.2.3 Uniendo y eliminando clases

Aun cuando las matrices de confusión no han mostrado hasta ahora ninguna superposición obvia, sería interesante probar si se pueden obtener mejores clasificadores si se crean subclasificadores agrupando pares de clases, como los indicados en la Tabla 4.8.

El primer intento fue unir las instancias de Lisi y Amorosos, asumiendo que los poemas de Canta sola a Lisi sean Amorosos y que sea posible que al agrupar dichas instancias se obtenga un mejor clasificador. Entonces las nuevas clases serían Lisi-Amorosos, Satíricos y FMR. Como resultado el árbol de clasificación devolvió resultados más equilibrados, aunque este clasificador solo obtuvo una exactitud del 52,83%. Considerando que con un dataset de tres clases la probabilidad básica de elegir aleatoriamente la clase correcta es del 33%, es evidente que este clasificador no provee más mejoras de las esperadas. Inclusive si los resultados mejorasen la exactitud de la clasificación, considerando tres de cuatro clases, se esperaba que exista esta mejora por lo menos por la simple razón de que tres clases son más fáciles de clasificar que cuatro.

Clases consideradas	Distribución	Exactitud
Amorosos y FMR	21 de 69	61,39%
Satírico y FMR	73 de 69	71,54%
Amorosos y Satíricos	21 de 73	73,40%

**Tabla 4.8.** Separando las clases por pares

#### 4.2.2.4 Otras técnicas de aprendizaje

Además de los árboles de decisión, se ha intentado utilizar otras técnicas con el fin de probar si eran las adecuadas para el dataset original (185 instancias para 4 clases).

Algunas de las técnicas probadas fueron Naïve Bayes [47] (55,67% de exactitud), Support Vector Machines [49] (53,51% de exactitud), Neural Networks (específicamente Multilayer Perceptron [43], con una exactitud del 53,51%), K\* [26] (52,97% de exactitud), Adaboost M1 [40] (52,43% de exactitud). Por otra parte, las matrices de confusión tienen las mismas proporciones que las producidas por los árboles de decisión.

### 4.2.3 Discusión

Los clasificadores que se construyeron en base al análisis emocional propuesto en este trabajo, dieron resultados razonablemente buenos. Incluso se hubieran podido construir mejores clasificadores con un número mayor de instancias (al menos para algunas clases como Lisi), al final está claro que la información emocional proporcionó datos valiosos para poder predecir la clasificación a la que pertenecía un poema dado. Eso significa que es posible concluir que existe una relación entre las emociones detectadas en los poemas y la clasificación realizada por Bleca.

Por ejemplo, es fácil notar que los poemas Amorosos de Quevedo son poemas tristes. En la misma línea los poemas escritos a esa mujer imaginaria llamada Lisi son altos en *Anger*, hecho que puede estar asociado a que en estos poemas se reflejan “frustraciones por las expectativas no logradas” [92] en la espera del amor imposible.

En cuanto a los resultados de los diferentes experimentos llevados a cabo, algunas de las conclusiones que se pueden extraer son:

- No se ha encontrado una combinación menor a cuatro clases que produzca mejor clasificación, es decir: el clasificador basado en las cuatro emociones tuvo una mejora del 112% sobre la selección *uninformed* o que no está basado en ningún parámetro (56,22% vs. 25% de exactitud). En el caso del clasificador de dos clases solo mejoró un 46,8% sobre la selección sin clasificar (73,4% vs. 50% de exactitud). Estos resultados implican que, desde el punto de vista del análisis sentimental, las cuatro clases son diferentes una de la otra.
- Los experimentos basados en datasets con resampling, generan resultados muy optimistas, esto significa que: las instancias de las clases minoritarias son replicadas un promedio de ocho veces, suficientes datos como para producir una propia hoja dentro del árbol de clasificación. Este tipo de estrategias son útiles para mostrar que se pueden obtener mejores resultados si se tuvieran disponibles mas instancias en las clases minoritarias, es decir si la cantidad de instancias estuvieran balanceadas en las cuatro clases.
- Ninguna de las técnicas principales de reconocimiento de patrones producen resultados más significativos que los Árboles de Decisiones, por esta razón, se cree que los modelos producidos por los Árboles de Decisiones y las reglas derivadas pueden ser analizadas para entender la relación entre los valores sentimentales y las categorías como poemas.

En ese sentido, los patrones principales que pueden ser extraídos de la observación de los árboles producidos en los diferentes experimentos, que se observan en la Tabla 5.2. En todos los casos, los términos "Alto" y "Bajo" son relativos y deben interpretarse en contraste con otros poemas, y con los valores detectados para cada dimensión emocional.

<i>Clase</i>	<i>Joy</i>	<i>Sadness</i>	<i>Anger</i>	<i>Fear</i>
<b>Lisi</b>	-	Bajo	Alto	Bajo
<b>Amorosos</b>	Bajo	Alto	-	-
<b>Satírico</b>	-	Bajo	Bajo	Mayormente Bajo
<b>FMR</b>	-	Alto(Si Fear=Bajo)	Alto	Alto

**Tabla 5.1. Patrones principales para cada clase**

Una posible fuente de ruido que encontramos en este trabajo fueron los poemas en los que no se encuentran palabras que expresen emociones. Es decir, poemas con un nivel emocional = 0 en cada una de las cuatro dimensiones. La cuestión principal era decidir si, sabiendo que ninguna emoción se detectó en un poema dado nos proporciona información útil en el momento de llevar a cabo la clasificación.

Pero al final hemos decidido trabajar con todos los poemas, independientemente de su valor, debido a que, aunque a primera vista parezcan poemas que no proveen información para los fines de la clasificación, la naturaleza del poema no cambia por el hecho de que sus resultados emocionales sean cero.



## 5 Conclusiones y trabajos futuros

---

### 5.1 Conclusiones

En este trabajo se ha presentado, en **primer lugar**, la propuesta de nuestro enfoque para la detección de emociones en textos mediante la construcción de conjuntos emocionales. La construcción de dichos conjuntos emocionales se ha diseñado ad-hoc con motivo de esta investigación, y para ello nos hemos centrado en las cuatro emociones básicas definidas por Zinck y Newen [92], entre otros; esto es: *joy*, *sadness*, *anger* y *fear*.

Nuestro enfoque está basado en la recuperación de información a través de un procedimiento de traducción (*cross-linguistic information retrieval*). Como punto de partida se tomaron las cuatro emociones básicas antes mencionadas y, a continuación, se procedió a completar estos cuatro conjuntos con palabras que de alguna forma u otra evocasen a cada una de dichas emociones. En el enfoque que se propone en este trabajo, se ha decidido trabajar con el inglés como idioma inicial y el español como idioma final, basándonos en el recurso lingüístico online bab.la<sup>3</sup>.

En **segundo lugar** con el fin de evaluar la adecuación de nuestra propuesta (los conjuntos emocionales), se realizó un estudio comparativo de ésta con S-ANew [69], centrado, concretamente, en comparar como nuestra propuesta y S-ANew clasifican las palabras de sus conjuntos y de qué tipo son las palabras contenidas en ambos. La gran diferencia entre nuestra propuesta y S-ANew yace en que ellos trabajan con palabras que se utilizan en la vida cotidiana que de alguna forma despiertan emociones y han sido calificadas con el criterio subjetivo de un grupo de personas. En cambio, en el caso de los conjuntos emocionales, la propuesta es trabajar con palabras que a través de su significado (de diccionario) evocan una emoción, sin la intervención subjetiva del criterio humano.

Una vez realizada la comparación de ambos enfoques se concluye que existe una clara correspondencia entre la descripción de las palabras propuestas por S-ANew y las palabras contenidas en nuestra propuesta. Además, se puede concluir que nuestra propuesta ha demostrado que proporciona un mecanismo coherente capaz de construir conjuntos de palabras “evocadoras” de emociones asociadas a emociones concretas. Por otra parte, su enfoque *cross-linguístico* permite que se aplique desde una perspectiva multilenguaje, con el fin de crear conjuntos emocionales en otros idiomas.

En **tercer lugar**, una vez construida y validada nuestra propuesta de conjuntos emocionales, hemos procedido a aplicarlos para el análisis emocional de textos literarios y conversaciones con el fin de realizar la categorización no supervisada de perfiles emocionales y de poemas de diferentes autores.

---

<sup>3</sup> Enero del 2012, <http://bab.la>

La realización de los estudios anteriores estaba asociada, desde el comienzo del trabajo, al planteamiento de algunas hipótesis y preguntas, las cuáles examinamos a continuación desde la perspectiva de los resultados obtenidos.

Respecto a las hipótesis planteadas, la primera de ellas era:

1. *Al analizar emocionalmente cuentos infantiles de cinco autores, los rangos emocionales tendrían que ser parecidos, debido a que cuando se habla de cuentos infantiles se asume que la emoción dominante tendría que ser la felicidad.*

Los resultados obtenidos tras el análisis emocional realizado en el apartado 4.1.1 nos muestran que nuestra primera hipótesis no es aplicable, debido a que los resultados presentan una clara división en dos clústeres emocionales: El primer clúster está formado por los tres escritores que concentran el porcentaje de *joy* más alto, siendo estos los **Hnos. Grimm, Andersen y Perrault**; el segundo clúster está formado por los dos escritores que concentran el porcentaje de *joy* más bajo y *fear* con el porcentaje más alto, **Pardo Bazán y Bécquer**.

De este análisis se concluye que ambos clústeres son inversamente proporcional con respecto al *joy* y al *fear*, es decir que, en el primer clúster mientras más alto sea el nivel de *joy*, más bajo será el nivel de *fear*, y el segundo clúster muestra que, mientras más bajo sean los niveles de *joy* más altos serán los de *fear*. Estos resultados podrían sugerir que es posible identificar, de alguna manera, el perfil emocional de los cuentos con la época, social o temporal, en que éstos fueron escritos, aunque se trate del mismo tipo de género literario.

En cuanto a la segunda de nuestras hipótesis, que se planteaba sobre la premisa anterior:

2. *Si se analizan los trabajos de dos escritores famosos tanto en género lírico como en el narrativo, tendríamos resultados emocionales muy parecidos en cada uno de los géneros.*

En este caso, en relación al género narrativo, en los cuentos escritos por los dos escritores objeto del estudio se ha encontrado que la emoción predominante es *fear*, y en proporciones parecidas: **Bécquer** ha obtenido un promedio del 39,75%, mientras **Benedetti** ha obtenido un promedio del 36,66%. La diferencia entre ambos se encuentra en cuál es la segunda emoción más fuerte en cada uno de ellos, siendo *sadness* en el caso de **Bécquer**, con el 24,36%, y *joy* en el de **Benedetti**, con el 28,61%.

En relación al género lírico, ambos escritores manifiestan diferencias emocionales notorias: **Bécquer**, por su parte, continúa transmitiendo *fear* como emoción predominante, y con mayor

proporción que el caso de la narrativa, con un promedio de 42,44%. En el caso de **Benedetti**, ahora la emoción predominante cambia a ser *sadness*, con el 32,06%, dejando en segundo lugar a *fear* con el 26,73%. Además de este cambio, se puede observar un mayor equilibrio entre la proporción de las cuatro emociones en sus poemas.

Estos resultados podrían indicar que existen trazas de algún tipo de perfil emocional propio de cada escritor.

Además de las hipótesis anteriores, y asociadas a ellas, se plantearon algunos interrogantes que marcaron los objetivos de algunos de los estudios realizados:

1. *¿El sistema de reconociendo de emociones sería capaz de reconocer las diferencias en los sentimientos transmitidos en dos grupos de conversaciones, realizadas en un contexto específico?*

En este trabajo el contexto elegido fue el intercambio de correos electrónicos en un ambiente académico. En este sentido, se encontró que en los correos electrónicos intercambiados entre **profesor – estudiante**, *joy* es la emoción con los niveles más altos y *anger* se encuentra ausente en estos resultados. A diferencia de los correos intercambiados entre compañeros de trabajo, en las conversaciones **profesor – profesor** las emociones están más equilibradas y, aunque *joy* sigue siendo la emoción más alta de las cuatro, *sadness* y *anger* tiene valores iguales.

Al ver que las conversaciones entre compañeros de trabajo (en este caso profesor – profesor) son emocionalmente más equilibradas que las conversaciones intercambiadas entre profesor – estudiante, se puede concluir que dicho equilibrio-desequilibrio emocional es el factor que identifica a estos dos tipos de conversaciones.

Y, en cuanto al último interrogante planteado:

2. *¿Si se toman los poemas de Francisco de Quevedo y una clasificación manual de los mismos, el sistema de reconocimiento propuesto sería capaz de realizar dicha clasificación automáticamente?*

En este caso, el objetivo del experimento fue determinar si un clasificador con los resultados del análisis emocional de los poemas de Quevedo era capaz de reproducir la clasificación de José Manuel Blecua [66]. En la edición de Blecua el número de poemas recolectados fue de 185 bajo 4 categorías emocionales o afectivas: Amorosos, Canta sola a Lisi, Satíricos y Poemas filosóficos, morales y religiosos.

En este caso, se puede concluir que un clasificador con este tipo de información es capaz de producir una clasificación más del doble de precisión que una selección aleatoria. En otras palabras, se han encontrado evidencias de que el análisis emocional sobre textos puede proporcionar información valiosa para la categorización de los mismos.

## **5.2 *Trabajos futuros***

Relacionado con los estudios realizados sobre la detección y análisis de emociones en textos de diversos autores, quedan muchas líneas de investigación aún abiertas. En la actualidad, estamos analizando la producción literaria de otros autores con el fin de aumentar los datos de que disponemos y avanzar en la posible identificación de perfiles, o trazas, emocionales. Hasta la fecha, los resultados obtenidos son muy prometedores en ese sentido. Al mismo tiempo, tenemos como tarea la mejora del contenido de los conjuntos emocionales, y también estamos trabajando en la implementación de las herramientas de reconocimiento de emociones en otros idiomas diferentes del español.

Además, se propone la posibilidad de utilizar ese tipo de reconocimiento de las emociones en textos como una herramienta útil para mejorar los modelos de usuario en diferentes contextos, como puede ser, aplicándolo en plataformas de e-learning o en sistemas de Help Desk. Una posible funcionalidad dentro de estos contextos puede ser construir el perfil emocional de un usuario dado; es decir, determinar el patrón emocional en base a la mayoría de sus textos. Una vez que el perfil emocional se construye, el sistema podría monitorizar cada nuevo trabajo que el usuario escribiera, y generar una alerta cuando un texto tiene una estampa emocional que se aleja de su patrón habitual.

En el caso de la clasificación emocional de los poemas de Francisco de Quevedo, para el futuro hemos planeado dos líneas de investigación: por un lado, necesitamos explorar porqué algunos poemas han sido mal clasificados. Es evidente, que ningún clasificador automático es perfecto, pero es sería interesante verificar si al mejorar la detección de la emoción lográramos reducir los errores de clasificación. Por otro lado, realizaremos pruebas en torno a esta propuesta sobre el trabajo de otros escritores y sus clasificaciones.







REFERENCIAS

Y

APENDICES



# Referencias

---

## Referencias Web

- [1] *Epinions. Unbiased Reviews by Real People*. Available: <http://www.epinions.com/>. Last retrieved on June, 2013.
- [2] *The HUMAINE Portal. Research on Emotions and Human-Machine Interaction*. Available: <http://emotion-research.net/>. Last retrieved on June, 2013.
- [3] *Bab.la, El portal lingüístico*. Available: <http://bab.la/>. Last retrieved on January, 2012.
- [4] *Ex-Sense Business Version*. Available: <http://www.nemesysco.com/sense.html>. Last retrieved on April, 2012
- [5] *Biblioteca Digital "Ciudad Seva"*. Available: <http://www.ciudadseva.com/bdcs/bdcs.htm>. Last retrieved on March, 2013.
- [6] *Tool kit for processing text using computational linguistics*. Available: <http://alias-i.com/lingpipe/>. Last retrieved on February 2013.

## Referencias

- [7] S. Asur and B. A. Huberman, "Predicting the future with social media," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, 2010, pp. 492-499.
- [8] A. Azcarate, F. Hageloh, K. V. D. S and R. Valenti, "Automatic facial emotion recognition," 2005. *Universiteit van Amsterdam*.
- [9] A. Batliner and R. Huber, "Speaker Characteristics and Emotion Classification," vol. 4343, pp. 138-151, 2007.
- [10] A. Batliner, S. Steidl, B. Schuller, D. Seppi, K. Laskowski, T. Vogt, L. Devillers, L. Vidrascu, N. Amir, L. Kessous and V. Aharonson, "Combining efforts for improving automatic classification of emotional user states," in *Proc. IS-LTC 2006, Ljubliana*, 2006, pp. 240-245.
- [11] P. K. Bhowmick, "Reader perspective emotion analysis in text through ensemble based multi-label classification framework," *Computer and Information Science*, vol. 2, pp. P64, 2009.
- [12] P. K. Bhowmick, A. Basu and P. Mitra, "Classifying emotion in news sentences: When machine classification meets human classification," *International Journal on Computer Science and Engineering*, vol. 2, pp. 98-108, 2010.

- [13] C. A. Bliss, I. M. Kloumann, K. D. Harris, C. M. Danforth and P. S. Dodds, "Twitter reciprocal reply networks exhibit assortativity with respect to happiness," *Journal of Computational Science*, vol. 3, pp. 388-397, 9, 2012.
- [14] M. M. Bradley and P. J. Lang, "Affective norms for english words (ANEW): Instruction manual and affective ratings," Tech. Rep. Technical Report C-1, 1999.
- [15] M. M. Bradley, "Emotional memory: A dimensional analysis," in *Emotions: Essays on Emotion Theory*, S. H. M. Van Goozen, N. E. Van De Poll and J. A. Sergeant, Eds. Routledge, 1994, pp. 97-134.
- [16] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J. Behav. Ther. Exp. Psychiatry*, vol. 25, pp. 49-59, 3, 1994.
- [17] J. Brooke, M. Tofiloski and M. Taboada, "Cross-Linguistic Sentiment Analysis: From English to Spanish," *Proceedings of the International Conference RANLP-2009*, pp. 50-54, September, 2009.
- [18] C. Bueno, J. A. Rojo and P. Rodríguez, "An Experiment on Semantic Emotional Evaluation of Chats," *The Fifth International Conference on Advances in Semantic Processing*, 2011.
- [19] R. A. Calvo and S. D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," *Affective Computing, IEEE Transactions on*, vol. 1, pp. 18-37, 2010.
- [20] G. Caridakis, K. Karpouzis and S. Kollias, "User and context adaptive neural networks for emotion recognition," *Neurocomputing*, vol. 71, pp. 2553, 2008.
- [21] V. Carvalho and C. Chao, "Sentiment retrieval in popular music based on sequential learning," in *Proc. ACM SIGIR*, 2005.
- [22] D. Cassany Comas, "Aproximaciones a la lectura crítica: teoría, ejemplos y reflexiones," *Tarbiya: Revista De Investigación e Innovación Educativa*, vol. 32, pp. 113-132, Madrid, 2003, 2003.
- [23] G. Castellano, M. Mortillaro, A. Camurri, G. Volpe and K. Scherer, "Automated analysis of body movement in emotionally expressive piano performances," *Music Perception: An Interdisciplinary Journal*, vol. 26, pp. 103-119, 2008.
- [24] G. Castellano, R. Bresin, A. Camurri and G. Volpe, "User-Centered Control of Audio and Visual Expressive Feedback by Full-Body Movements," vol. 4738, pp. 501-510, 2007.
- [25] Z. Chuang and C. Wu, "Multi-modal emotion recognition from speech and text," *International Journal of Computational Linguistics and Chinese Language Processing*, vol. 9, pp. 1-18, 2004.

- [26] J. G. Cleary and L. E. Trigg, "K\*: An instance-based learner using an entropic distance measure," in *In Proceedings of the 12th International Conference on Machine Learning*, 1995, pp. 108-114.
- [27] C. Conati, "Combining Cognitive Appraisal and Sensors for Affect Detection in a Framework for Modeling User Affect," vol. 3, pp. 71-84, 2011.
- [28] C. Conati and H. Maclaren, "Empirically building and evaluating a probabilistic model of user affect," *User Modeling and User-Adapted Interaction*, vol. 19, pp. 267-303, 2009.
- [29] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz and J. G. Taylor, "Emotion recognition in human-computer interaction," *Signal Processing Magazine, IEEE*, vol. 18, pp. 32-80, 2001.
- [30] S. K. D'Mello and A. Graesser, "Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features," *User Modeling and User-Adapted Interaction*, vol. 20, pp. 147-187, 2010.
- [31] S. Dahl and A. Friberg, "Visual perception of expressiveness in musicians' body movements," *Music Perception: An Interdisciplinary Journal*, vol. 24, pp. 433-454, 2007.
- [32] S. Dahl and A. Friberg, "Expressiveness of Musician's Body Movements in Performances on Marimba," vol. 2915, pp. 479-486, 2004.
- [33] J. W. Davidson, "Visual perception of performance manner in the movements of solo musicians," *Psychology of Music*, vol. 21, pp. 103-113, 1993.
- [34] L. Devillers and L. Vidrascu, "Real-life emotions detection with lexical and paralinguistic cues on human-human call center dialogs," in *Ninth International Conference on Spoken Language Processing*, 2006, .
- [35] P. S. Dodds and C. M. Danforth, "Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents," *Journal of Happiness Studies*, vol. 11, pp. 441-456, 2010.
- [36] P. S. Dodds, K. D. Harris, I. M. Kloumann, C. A. Bliss and C. M. Danforth, "Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter," *PLoS ONE*, vol. 6, no 12, pp e26752, 2011.
- [37] E. Douglas-Cowie, R. Cowie, I. Sneddon, C. Cox, O. Lowry, M. McRorie, J. Martin, L. Devillers, S. Abrilian, A. Batliner, N. Amir and K. Karpouzis, "The HUMAINE Database: Addressing the Collection and Annotation of Naturalistic and Induced Emotional Data," vol. 4738, pp. 488-500, 2007.
- [38] P. Ekman, "Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique." *Psychology Bulletin*, vol. 115, pp. 268-287, 1994.

- [39] M. Feidakis, T. Daradoumis and S. Caballe, "Emotion measurement in intelligent tutoring systems: What, when and how to measure," in *Intelligent Networking and Collaborative Systems (INCoS), 2011 Third International Conference on*, 2011, pp. 807-812.
- [40] Y. Freund and R. E. Schapire, "Experiments with a new Boosting algorithm," *Proc.ICML-1996*, 1996.
- [41] A. Friberg, "A fuzzy analyzer of emotional expression in music performance and body motion," in *Proceedings of Music and Music Science, Stockholm 2004*, 2005.
- [42] M. Grimm, E. Mower, K. Kroschel and S. Narayanan, "Combining categorical and primitives-based emotion recognition," in *14th European Signal Processing Conference (EUSIPCO), Florence, Italy*, 2006.
- [43] S. Haikin, *Neural Networks: A Comprehensive Foundation*. NY: Pearson Education, 1998.
- [44] G. E. Hinton, S. Osindero and Y. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, pp. 1527-1554, jul, 2006.
- [45] G. Hirst, "Does Conversation Analysis Have a Role in Computational Linguistics?" *Computational Linguistics*, 1991, vol. 17 no 2, pp. 211-227.
- [46] T. Hu, L. C. De Silva and K. Sengupta, "A hybrid approach of NN and HMM for facial emotion classification," *Pattern Recog. Lett.*, vol. 23, pp. 1303-1310, 9, 2002.
- [47] G. H. John and P. Langley, "Estimating continuous distributions in bayesian classifiers," in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, Montréal, Québec, Canada, 1995, pp. 338-345.
- [48] S. D. Kamvar and J. Harris, "We feel fine and searching the emotional web," in *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, Hong Kong, China, 2011, pp. 117-126.
- [49] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya and K. R. K. Murthy, "Improvements to Platt's SMO Algorithm for SVM Classifier Design," *Neural Comput.*, vol. 13, pp. 637-649, mar, 2001.
- [50] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, pp. 2067-2083, 2008.
- [51] I. M. Kloumann, C. M. Danforth, K. D. Harris, C. A. Bliss and P. S. Dodds, "Positivity of the english language," 2012 , PloS one, vol. 7, no 1, pp. e29484.
- [52] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, 1995, pp. 1137-1143.



- [53] Z. Kozareva, B. Navarro, S. Vázquez and A. Montoyo, "UA-ZBSA: A headline emotion classification through web information," in *Proceedings of the 4th International Workshop on Semantic Evaluations*, Prague, Czech Republic, 2007, pp. 334-337.
- [54] A. D. I. Kramer, "An unobtrusive behavioral model of gross national happiness," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Atlanta, Georgia, USA, 2010, pp. 287-290.
- [55] C. M. Lee and S. S. Narayanan, "Toward detecting emotions in spoken dialogs," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, pp. 293-303, 2005.
- [56] C. Lee, E. Mower, C. Busso, S. Lee and S. Narayanan, "Emotion recognition using a hierarchical binary decision tree approach," *Speech Commun.*, vol. 53, pp. 1162-1171, 0, 2011.
- [57] C. M. Lee, S. Yildirim, M. Bulut, A. Kazemzadeh, C. Busso, Z. Deng, S. Lee and S. Narayanan, "Emotion recognition based on phoneme classes," in *Proc. ICSLP'04*, 2004, pp. 889-892.
- [58] B. Levin, *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago press Chicago, 1993.
- [59] K. H. Lin, C. Yang and Hsin-Hsi Chen, "Emotion classification of online news articles from the reader's perspective," in *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on*, 2008, pp. 220-226.
- [60] M. J. Lyons, J. Budynek and S. Akamatsu, "Automatic classification of single facial images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, pp. 1357-1362, 1999.
- [61] X. Mao and Z. Li, "Implementing emotion-based user-aware e-learning," in *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, Boston, MA, USA, 2009, pp. 3787-3792.
- [62] J. M. Martin, A. Ortigosa and R. M. Carro, "SentBuk: Sentiment analysis for e-learning environments," in *Computers in Education (SIIE), 2012 International Symposium on*, 2012, pp. 1-6.
- [63] R. Mihalcea and H. Liu, "A corpus-based approach to finding happiness," *Proceedings of the AAAI Spring Symposium on Computational Approaches to Weblogs*, pp. 19, 2006.
- [64] A. Osherenko, "Towards semantic affect sensing in sentences," *Proceedings of the AISB 2008 Symposium on Affective Language in Human and Machine*, pp. 41-44, 2008.
- [65] C. Padgett and G. W. Cottrell, "Representing face images for emotion classification," *Advances in Neural Information Processing Systems*, pp. 894-900, 1997.

- [66] F. d. Quevedo, "Poemas escogidos, edición de José Manuel Blecua," *Madrid, Editorial Castalia*, 1972.
- [67] F. d. Quevedo and C. D. Johnson, *Selected Poetry of Francisco De Quevedo: A Bilingual Edition*. University of Chicago Press, 2009.
- [68] J. R. Quinlan, "Induction of decision trees," *Mach. Learning*, vol. 1, pp. 81-106, 1986.
- [69] J. Redondo, I. Fraga, I. Padrón and M. Comesaña, "The Spanish adaptation of ANEW (Affective Norms for English Words)," *Behavior Research Methods*, vol. 39, pp. 600-605, 2007.
- [70] M. G. D. Ríos and A. Gravano, "Spanish DAL: A Spanish Dictionary of Affect in Language," *Wassa 2013*, pp. 21, 2013.
- [71] P. Rodríguez, A. Ortigosa and R. M. Carro, "Extracting emotions from texts in E-learning environments," in *Complex, Intelligent and Software Intensive Systems (CISIS), 2012 Sixth International Conference on*, 2012, pp. 887-892.
- [72] B. Schuller, A. Batliner, S. Steidl and D. Seppi, "Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge," *Speech Commun.*, vol. 53, pp. 1062-1087, 0, 2011.
- [73] L. Shi, R. Mihalcea and M. Tian, "Cross language text classification by model translation and semi-supervised learning," in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, Cambridge, Massachusetts, 2010, pp. 1057-1067.
- [74] P. J. Stone, D. C. Dunphy and M. S. Smith, *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press, 1966,.
- [75] C. Strapparava and R. Mihalcea, "Learning to identify emotions in text," in *Proceedings of the 2008 ACM Symposium on Applied Computing*, Fortaleza, Ceara, Brazil, 2008, pp. 1556-1560.
- [76] C. Strapparava and R. Valitutti, "WordNet-affect: An affective extension of WordNet," in *In Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon, 2004, pp. 1083-1086.
- [77] J. M. Susskind, G. E. Hinton, J. R. Movellan and A. K. Anderson, "Generating facial expressions with deep belief nets," *Affective Computing, Emotion Modelling, Synthesis and Recognition*, pp. 421-440, 2008.
- [78] I. V. Tetko, D. J. Livingstone and A. I. Luik, "Neural network studies. 1. Comparison of overfitting and overtraining," *J. Chem. Inf. Comput. Sci.*, vol. 35, pp. 826-833, 09/01; 2013/04, 1995.

- [79] Tin Lay Nwe, Foo Say Wei and L. c. De Silva, "Speech based emotion classification," in *TENCON 2001. Proceedings of IEEE Region 10 International Conference on Electrical and Electronic Technology*, 2001, pp. 297-301 o.1.
- [80] D. Ververidis, C. Kotropoulos and I. Pitas, "Automatic emotional speech classification," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, 2004, pp. -593-6 o.1.
- [81] K. Voll and M. Taboada, "Not all words are created equal: Extracting semantic orientation as a function of adjective relevance," in *AI 2007: Advances in Artificial Intelligence*, M. Orgun and J. Thornton, Eds. Springer Berlin Heidelberg, 2007, pp. 337-346.
- [82] C. Whissell, "The dictionary of affect in language," *Emotion: Theory, Research, and Experience*, vol. 4, pp. 94, 1989.
- [83] C. Whissell, "A comparison of two lists providing emotional norms for English words (ANEW and the DAL)," *Psychol. Rep.*, vol. 102, pp. 597-600, Apr, 2008.
- [84] I. H. Witten and E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations (The Morgan Kaufmann Series in Data Management Systems)," 1999.
- [85] T. Wu and S. Jeng, "Extraction of segments of significant emotional expressions in music," in *Workshop on Computer Music and Audio Technology*, 2006, pp. 76-80.
- [86] Xi Li, Jidong Tao, M. T. Johnson, J. Soltis, A. Savage, K. M. Leong and J. D. Newman, "Stress and emotion classification using jitter and shimmer features," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 2007, pp. -1081-1084.
- [87] D. Yang and W. Lee, "Disambiguating music emotion using software agents." in 2004, pp. 218-223.
- [88] Y. Yang, Y. Lin, Y. Su and H. H. Chen, "A Regression Approach to Music Emotion Recognition," *Trans.Audio, Speech and Lang.Proc.*, vol. 16, pp. 448-457, feb, 2008.
- [89] M. Yassine and H. Hajj, "A framework for emotion mining from text in online social networks," in *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*, 2010, pp. 1136-1142.
- [90] G. P. Young, "Imagery in Quevedo's love poetry," 1974.
- [91] F. Yu, E. Chang, Y. Xu and H. Shum, "Emotion detection from speech to enrich multimedia content," in *Proceedings of the Second IEEE Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing*, 2001, pp. 550-557.
- [92] A. Zinck and A. Newen, "Classifying emotion: a developmental account," *Synthese*, vol. 161, pp. 1-25, 2008.



# Apéndice 1: Construyendo los conjuntos emocionales en inglés

<i>Joy &amp; Joyful (25 palabras)</i>	<i>Sadness &amp; Sad (30 palabras)</i>	<i>Anger &amp; Angry (27 palabras)</i>	<i>Fear &amp; Afraid (25 palabras)</i>
animated	blues	to annoy	afflicting
cheerful	brokenheart	annoyance	to be afraid
cheerfulness	dejected	annoyed	agony
convivial	deplorable	to apotheosize	apprehension
delight	depressed	to arouse	anxious
elated	depression	braveness	concern
elation	distressing	choler	distress
exhilaration	dreary	cholera	distressed
fortunate	gaunt	courage	distressing
Gaiety	gloom	courageousness	dread
glad	Gloominess	to excite	fearful
gladden	gloomy	furious	fearfulness
gleeful	glum	heated-up	fright
happy	heartbroken	impassioned	heartbreaking
jolly	heaviness	ire	nervous
joyousness	lamentable	irritable	revere
jubilant	low spirits	mad	reverence
jubilation	lugubriousness	nerve	torment
lively	melancholy	rage	tormented
lucky	miserable	raging	tormenting
merry	pitiful	rebound	trouble
mirth	slough	ricochet	uneasy
overjoyed	somber	red	venerate
playfulness	sorrow	tempestuous	worried
rejoice	sorrowfulness	wild	worrisome
	sorry	wrath	
	sullen		
	unhappiness		
	wistful		
	wistfulness		

Tabla Ap.1. Lista inicial de los conjuntos emocionales en inglés



## Apéndice 2: Construyendo los conjuntos emocionales en español

<i>Joy &amp; Joyful (90 palabras)</i>				
afortunado	Apasionar	divinidad	fruición	promesa
agradable	aprecio	encantado	goce	recrearse
agradar	bueno	encantador	gozo	recreo
agradezco	celebración	encanto	gozoso	regocijar
alborozado	celebrar	entusiasmar	gratificante	regocijarse
alborozar	complace	entusiasmo	gustar	regocijo
alborozarse	complacencia	entusiasta	gustaría	regodear
alborozo	complacer	estupendo	gusto	rico
alegrar	congratular	euforia	gustosamente	risueño
alegrarse	congratulo	eufórico	honor	satisfacción
alegre	contento	fausto	interesante	satisfacer
alegría	deleitar	favorable	jovial	satisfecho
aleluya	deleitarse	felices	júbilo	sonriente
alentador	deleite	felicitación	jubilosa	suerte
ameno	delicia	felicitar	placentero	ventaja
animado	dichoso	felicitarnos	placer	venturoso
apasionado	disfrutar	feliz	positivo	vivaz
apasionante	divertido	festiva	privilegio	vivo

**Tabla Ap.2.** Conjunto de palabras en español para “Joy y Joyful”

<i>Sadness &amp; Sad (166 palabras)</i>				
abatido	contristarse	desilusionado	insignificante	perdón
abatimiento	costra	desmayar	insuficiente	perdonar
aberrante	crisis	desolación	intolerable	pesadez
abhorrecible	criticable	desolador	lacra	pesadumbre
aburrido	cruel	detestable	lamentable	pesar
acongojado	decaído	disculparse	lamentar	pesimismo
acongojar	decaimiento	disculpas	lastima	pesimista
aflicción	decepción	disgusto	lastimoso	piadosa
afligido	deflación	dolor	lóbrega	pobre
agorero	densidad	dramático	lóbrego	postración
alteración	denunciable	duelo	lúgubre	preocupa
angustia	deplorable	duro (revés)	luto	preocupación
angustioso	depre (col.)	duros	mal	preocupante
anómala	depresión	engurruñarse	mala (noticia)	quebrado
apagado	deprimido	entrecortado	malestar	quebranto
apenar	desafortunada	entristecer	melancolía	rayado (col.)
apenarse	desagradable	entristecerse	mezquino	resentido
apesadumbrado	desalentador	escandalosa	miserable	sentimiento
apesadumbrar	desalentadora	espinoso	miserablemente	sombrío
arrastrado	desamparo	execrable	morosidad	taciturno
arrepentido	desanimado	flagrante	moroso	tenebrosa
atolladero	desbaratar	fúnebre	mustio	tenebrosidad
atonía	desconsolador	funesto	negativo	tenebroso
atribular	descontento	horrible	nimias	tenue
blues	desesperado	impactante	nostalgia	terrible
calvario	desfavorecido	indigente	palpitar	tétrico
catastrofismo	desgarrador	indignante	patético	tiniebla
catástrofe	desgracia	indigno	paupérrimo	tragedia
cenagal	desgraciada	infame	pena	trágica
condenable	desgraciadamente	infelicidad	penar	trágico
conmovedora	desgraciado	infortunio	penoso	triste
consternación	desheredado	infrahumano	pensativo	tristeza
consternarse	desilusión	inquietante	penumbra	turbio
				vergüenza

**Tabla Ap.3. Conjunto de palabras en español para “Sadness y Sad”**



<i>Anger &amp; Angry (169 Palabras)</i>				
aborrascado	delirio	enfervorizar	hervir	odio
abroncarse	dementar	enfurecer	hostil	ofenderse
absurdo	desabrirse	enfurecerse	hostilidad	orate
acalorado	desatado	enfurecido	huracanado	osadía
adulterar	desatinado	enojadizo	ilusionar	paroxismo
afán	desazón	enojado	incomodar	peligroso
agitación	desazonado	enojar	incomodarse	perturbar
agitado	desazonante	enojarse	incomodo	picarse
agitar	desazonar	enojo	incordiar	precipitadamente
airado	descaro	escocer	incordio	preocuparse
airarse	desencadenar	estallado	incordioso	protesta
alterarse	desenfrenado	estragos	indignación	rabia
amohinarse	desfachatez	exaltado	indignado	rabioso
animar	desmedido	exaltar	indignante	rampante
ánimo	despotricar	exasperación	indignar	rebotado
apasionado	despropósito	exasperante	insatisfecho	rebotar
apasionante	desquiciado	excitado	intempestivo	rebote
apasionar	devastado	excitante	ira	renegar
arriesgado	disgustar	excitar	iracundo	revuelta
arriesgar	disgusto	exultante	irracional	salvaje
arroyo	disparatado	fastidiado	irritable	susitar
asolar	embravecerse	fastidiar	irritación	tempestuoso
atrever	embroncarse	fastidiarse	irritado	temple
atrevido	enardecido	fastidio	irritar	tirantez
atrevimiento	encolerizar	fiero	irritarse	tormentoso
audacia	encolerizarse	firme	jorobar	trastornar
bravura	enconamiento	firmeza	loco	tumultuoso
cargar	encono	frenético	malestar	valentía
cólera	encorajarse	furia	miedo	valiente
coraje	endemoniarse	furibundo	molesta	valor
crispar	enfadado	furioso	molestar	ventoso
decepcionado	enfadar	furor	molestarse	vertiginoso
decidido	enfadarse	hartar	molesto	vociferación
decidir	enfado	harto	nervio	

**Tabla Ap.4. Conjunto de palabras en español para “Anger y Angry”**

<i>Fear &amp; Afraid (215 palabras)</i>				
acelerado	aterrorizar	dificultad	malestar	problema
acosado	atolladero	difícultoso	malintencionado	problemático
afección	atormentado	discordia	malo	recelar
afectado	atormentar	disgusto	martirio	recelo
afectar	atormentarse	disturbio	martirizar	reverencia
afligir	auxilio	doloroso	miedo	sacudió
afligidos	calamidad	dramático	miedosa	sangrante
agitación	calvario	duda	miserable	sobresaltar
agitada	castigado	emergencia	miseria	sobresalto
agonía	cautela	empeño	molestar	sufrimiento
alarma	conflictivo	enrarecido	molestarla	sufrir
alarmarse	conflicto	ensañarse	molestarse	suplicio
amenaza	consciente(ser c.)	entristecer	molestia	susto
amenazado	consternación	esforzado	mortificante	temblar
angustia	consternado	esfuerzo	mortificar	tembloroso
angustiada	consternador	espanto	mortificarse	temer
angustiado	consternar	fastidioso	necesidad	temerlo
angustiados	consternarse	flagelo	necesitado	temeroso
angustiante	contrariar	flagrante	nervioso	temerse
angustiar	convulso	golpeado	obstaculizar	temible
angustiar	crisis	golpear	obstáculo	temor
angustioso	cruel	grave	padecer	tensión
ansiedad	cruz	horror	padecimiento	terrible
ansiosamente	cuitado	impaciente	pánico	terror
ansioso	daño	incomodar	pavor	tormento
apenado	deformador	incómodo	pavoroso	tortura
apenar	deformar	infortunio	peligro	torturado
apesadumbrado	desagradable	inquietante	pena	torturar
aprehensión	desalentador	inquietarse	penoso	tragedia
aprensión	desamparo	inquieto	penuria	trastornar
aprensivo	desastre	inquietud	perjuicio	trastorno
aprieto	desconcertado	inquietudes	pesadumbre	tremendamente
apuro	desconcierto	interesado	pesar	tremendo
aquejado	desconsolar	intranquilizar	pesimismo	tribulación
aquejar	descorazonador	intranquilizarse	pesimista	tribulaciones
asustadizo	desesperación	intranquilo	podrido	triste

asustado	desesperado	intrigado	preocupación	tristeza
asustar	desesperante	lamentable	preocupaciones	turbación
asustarse	desgarrador	lamentablemente	preocupado	urgencia
atemorizado	desgracia	lamentar	preocupante	urgente
atemorizar	desgraciadamente	lamento	preocupar	venerar
atenazado	desolador	lío	preocuparse	victimas
atenazar	difícil	mal		zozobra

**Tabla Ap.5. Conjunto de palabras en español para “Fear y Afraid”**



## Apéndice 3: Listado de obras de G. A. Bécquer y M. Benedetti

---

<i>Cuentos de Bécquer</i>	
1	La corza blanca
2	La cueva de la mora
3	El Gnomo
4	Maese Pérez el Organista
5	El monte de la ánimas
6	Los ojos verdes
7	El rayo de luna
8	El Miserere
9	Creed en Dios
10	El beso
11	El caudillo de las manos rojas
12	El cristo de la calavera
13	Es raro
14	La ajorca de oro
15	La arquitectura árabe en Toledo
16	La creación
17	La cruz del diablo
18	La mujer de piedra
19	La promesa
20	La rosa de pasión
21	Las hojas secas
22	Tres fechas

**Tabla Ap.6. Listado de 22 cuentos de Bécquer**

<i>Rimas de Bécquer</i>	
1	RIMA LXXIII
2	RIMA XII
3	RIMA LXXIII
4	RIMA LXX
5	RIMA V
6	RIMA XII
7	RIMA IV
8	RIMA III
9	RIMA XL
10	RIMA LXXVI
11	RIMA XXVIII
12	RIMA XXV
13	RIMA XXXIV
14	RIMA XV
15	RIMA XXXVII

**Tabla Ap.7. Listado de 15 rimas de Bécquer**

<i>Cuentos de Benedetti</i>	
1	A Carlos Real de Azúa
2	Caramba y lástima
3	La lluvia y los hongos
4	No tenía lunares
5	Como un ladrón
6	Hoy y la alegría
7	Esta mañana
8	Aquí se respira bien
9	Sábado de Gloria
10	Los pocillos
11	No ha claudicado
12	El otro yo
13	A imagen y semejanza
14	Puentes como liebres
15	Soñó que estaba preso
16	Conciliar el sueño
17	Beatriz (La polución)
18	Beatriz (Una palabra enorme)
19	Historias de fantasmas
20	Idilio

**Tabla Ap.8. Listado de 20 cuentos de Benedetti**

<i>Poemas de Benedetti</i>	
1	Artigas
2	Ausencia de Dios
3	Bienvenida
4	Chau Pesimismo
5	Hagamos un trato
6	Hombre que mira la tierra
7	Parpadeo
8	Me sirve no me sirve
9	Señas del Che
10	Triste N° 1
11	Te quiero
12	Bonjour buon giorno guten morgen,
13	Si Dios fuera mujer
14	¿Por qué no hay más viajes a la luna?
15	Las soledades de Babel
16	Últimas golondrinas
17	El Sur también existe
18	Los formales y el frío
19	Poema frustrado
20	Chau número tres

**Tabla Ap.9. Listado de 20 poemas de Benedetti**